

Fr m: Parkin, Jeffrey  
Sent: Saturday, March 15, 2003 1:06 PM  
To: STIC-Biotech/ChemLib  
Subj ct: U.S. Serial No. 09/877,606

CF

Please search **SEQ ID NOS.: 3, 4, 7, 8, and 9**, vis-a-vis the aforementioned application (**09/877,606**) v. all relevant databases, including interference. Place results on BOTH paper and disk. Thanks!

JSP  
AU 1648  
CM01-8E15  
308-2227

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 3/18/03  
Date Completed: 3/27/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

## TYPE OF SEARCH:

NA Sequences: \_\_\_\_\_  
AA Sequences: 4-4 reverse  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

## VENDOR/COST (where applic.)

STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: EVIP/04  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4 p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:27:06 ; Search time 37.1429 Seconds

(without alignments)  
186.351 Million cell updates/sec

Title: US-09-877-606-3

Perfect score: 292  
Sequence: 1 NHTTWLEMDREINNTSLIH.....NEQELLELDKNASLMNWFNI 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: A.Geneseq\_101002.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	99.3	268	19	AA22820
2	290	99.3	268	23	ABG68291
3	290	99.3	344	23	ABB83400
4	290	99.3	420	15	AA853785
5	290	99.3	519	17	AAW00181
6	290	99.3	521	16	AA869997
7	290	99.3	853	19	AAW43066
8	290	99.3	856	7	AA60131
9	290	99.3	856	14	AA41025
10	290	99.3	856	14	AA41026

11	290	99.3	856	14	AA41027
12	290	99.3	856	14	AA41028
13	290	99.3	856	14	AA41029
14	290	99.3	856	14	AA41030
15	290	99.3	856	14	AA41031
16	290	99.3	856	14	AA41032
17	290	99.3	856	22	AA456972
18	290	99.3	856	22	AA456972
19	290	99.3	856	22	AA456972
20	290	99.3	856	22	AA456972
21	290	99.3	856	22	AA456972
22	290	99.3	856	22	AA456972
23	290	99.3	856	22	AA456972
24	290	99.3	856	22	AA456972
25	290	99.3	856	22	AA456972
26	290	99.3	856	22	AA456972
27	290	99.3	856	22	AA456972
28	290	99.3	856	22	AA456972
29	290	99.3	856	22	AA456972
30	290	99.3	856	22	AA456972
31	290	99.3	856	22	AA456972
32	290	99.3	856	22	AA456972
33	290	99.3	856	22	AA456972
34	290	99.3	856	22	AA456972
35	290	99.3	856	22	AA456972
36	290	99.3	856	22	AA456972
37	290	99.3	856	22	AA456972
38	290	99.3	856	22	AA456972
39	290	99.3	856	22	AA456972
40	290	99.3	856	22	AA456972
41	290	99.3	856	22	AA456972
42	290	99.3	856	22	AA456972
43	290	99.3	856	22	AA456972
44	290	99.3	856	22	AA456972
45	290	99.3	856	22	AA456972

# ALIGNMENTS

RESULT 1	AA22820	AA22820 standard; Protein; 268 AA.
ID	AA22820	
XX	AA22820	
AC	AA22820	
XX	AA22820	
DT	19-AUG-1999	(first entry)
XX	AA22820	
DE	SEQ ID NO. 16 from WO9820036.	
XX	AA22820	
KW	HIV; gp41 protein; constrained helical peptide; HIV infection;	
KW	vacine; antibody; viral membrane fusion; viral infectivity;	
KW	ligand affinity purification; protein A replacement;	
KW	immunoglobulin purification; epitope mimic.	
XX	AA22820	
OS	Human immunodeficiency virus.	
XX	AA22820	
PN	WO9820036-A1.	
XX	AA22820	
PD	14-MAY-1998.	
XX	AA22820	
PF	05-NOV-1997;	97WO-US20069.
XX	AA22820	
PR	16-JUN-1997;	97US-0876698.
XX	AA22820	
PA	06-NOV-1996;	96US-0743698.
XX	AA22820	
PI	(GETH) GENENTECH INC.	
XX	AA22820	
PI	Braisted A, Judice JK, McDowell RS, Phean JC, Starovasnik MA;	
XX	Wells JA.	
XX	WPI, 1998-286866/25.	

Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Wild type HIV-1 HX  
HIV-1 isolate LAI  
HIV viral envelope  
Amino acid sequenc  
HIV-1 gp41 envelop  
Sequence of Aas 60  
HIV gp41 LAI prote  
HIV gp41 LAI prote  
Recombinant gp41 p  
Recombinant HIV p  
SEQ ID NO. 17 from  
HIV-1 env protein.  
HIV-1 coat protein  
HIV gp41 DNA SEQ I  
Protein encoded by

09/78/7608

PT production of constrained helical peptide(s) by linking side chains  
PT on termin of octa:peptide - derived from human immunodeficiency  
PT virus gp41 protein, useful in vaccines for treatment and prevention  
PT of infection  
XX  
PS Claim 11; Page 156; 279pp; English.  
XX  
CC Peptides AAY22805-222917 are derived from Human immunodeficiency virus  
CC (HIV). Specifically, AAY22810-222910 are derived from gp41 proteins  
CC of known HIV virus strains (AAY22810, AAY22871, AAY22880, AAY22888 and  
CC AAY22903 represent consensus sequences of various sections of the gp41  
CC protein). Sequences derived from the peptides are used to produce  
CC constrained helical peptides of the invention. The constrained helical  
CC peptide is produced by synthesizing an octapeptide in which both terminal  
CC amino acids have a side-chain that includes a group able to form an amide  
CC bond, and cyclizing the octapeptide by reacting the specified side-chain  
CC residues with a difunctional linker to produce two amide bonds.  
CC The constrained helical peptides are used to treat or prevent HIV  
CC infection, especially as vaccines that generate antibodies that  
CC prevent viral membrane fusion or infectivity. Vaccines may contain  
CC constrained helical peptides derived from several different strains of  
CC HIV. The antibodies are also useful for diagnosing HIV infection. Other  
CC uses for the constrained helical peptides are in affinity purification  
CC of ligands (particularly where complete binding protein is not readily  
CC available, e.g. replacements for protein A in immunoglobulin  
CC purification); as epitope mimics for antibody production; for isolation  
CC of synthetic antibody clones from phage display libraries, or as stable  
CC forms of "floppy" peptides or proteins.  
XX  
SQ Sequence 268 AA;  
XX  
Query Match 99.3%; Score 290; DB 19; Length 268;  
Best Local Similarity 98.1%; Pred. No. 1.6e-23;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 NHTTWMEDREINNYTSLIHSLEESONQOEKNEDELLEDKASLNMWENI 52  
154 NHTTWMEDREINNYTSLIHSLEESONQOEKNEDELLEDKASLNMWENI 205  
XX  
RESULT 2  
ABG68291 standard; Protein; 268 AA.  
XX  
AC ABG68291;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Envelope protein gp41 from HIV clade B strain #10.  
XX  
KW HIV; glycoprotein; gp41; antigen; helical conformation;  
KW virus-induced membrane fusion; acquired immunodeficiency syndrome; AIDS;  
KW viral envelope protein; vaccine; virucide; anti-HIV.  
XX  
OS Human immunodeficiency virus type 1 clade B.  
XX  
PN US6271198-B1.  
XX  
PD 07-AUG-2001.  
XX  
PE 05-NOV-1997; 97US-0965056.  
XX  
PR 16-JUN-1997; 97US-049787P.  
PR 06-NOV-1996; 96US-0743698.  
PR 16-JUN-1997; 97US-0876698.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Braisted AC, Judice JK, McDowell RS, Phelan JC, Starovashnik MA;  
PI Wells JA;  
XX  
DR WPI: 2002-487624/52.  
XX  
PT

PT New cyclic peptides from human immune deficiency virus gp41, useful for  
PT treatment or prevention of HIV infection, are constrained to have  
PT alpha-helical conformation -  
XX  
PS Disclosure: Column 151-154; 175pp; English.  
XX  
CC The invention relates to cyclic peptides (A) with a constrained helical  
CC conformation, derived from gp41 (glycoprotein 41, a viral envelope  
CC protein) of human immunodeficiency virus (HIV). The cyclic  
CC peptides have formulas given in the specification part of which are  
CC derived from a consensus sequence of gp41 derived from HIV clades A, B,  
CC C, D, E or O. The peptides are used to cause induction of a specific  
CC immune response, resulting in antibodies that prevent virus-induced  
CC membrane fusion. The peptides are used to treat subjects with, or at risk  
CC of, HIV infection, either as antifusion/anti-infection agents or,  
CC preferably where associated with a carrier, as an immunogen (including as  
CC vaccine) to raise antibodies. The antibodies may be used for diagnosis or  
CC prevention/treatment of HIV infection (i.e. acquired immunodeficiency  
CC syndrome, AIDS), e.g. prevention of mother-to-child transmission or in  
CC cases of health care accidents. The peptides can be based on specific HIV  
CC strains, e.g. breakthrough isolates of HIV that have developed during  
CC vaccine trials, so a combination of them should cover a wide range of  
CC protection. The present sequence is gp41 protein from a particular  
CC HIV clade used to derive a consensus sequence of gp41.  
XX  
SQ Sequence 268 AA;  
XX  
Query Match 99.3%; Score 290; DB 23; Length 268;  
Best Local Similarity 98.1%; Pred. No. 1.6e-23;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
XX  
QY 1 NHTTWMEDREINNYTSLIHSLEESONQOEKNEDELLEDKASLNMWENI 52  
154 NHTTWMEDREINNYTSLIHSLEESONQOEKNEDELLEDKASLNMWENI 205  
XX  
RESULT 3  
ABB83400 standard; protein; 344 AA.  
XX  
AC ABB83400;  
XX  
DT 19-SEP-2002 (first entry)  
XX  
DE HIV gp41 LAI protein #1.  
XX  
KW HIV; gp41; LAI; virucide; anti-HIV; vaccine; HIV infection.  
XX  
OS Human immunodeficiency virus.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..23  
FT /label= Fusion\_Peptide  
FT Domain 173..194  
FT /label= Transmembrane\_domain  
XX  
PN WO200253587-A2.  
XX  
PD 11-JUL-2002.  
XX  
PE 04-JAN-2002; 2002WO-FR00031.  
XX  
PR 05-JAN-2001; 2001FR-0000141.  
PR 23-JAN-2001; 2001FR-0000848.  
XX  
PA (AVET ) AVENTIS PASTEUR.  
XX  
PI Brasseur R, Charlotaux B, Chevalier M, El Habib R, Krell T;  
PI Sodoyer R;  
XX  
DR WPI: 2002-528852/56.  
XX  
PT New mutant human immune deficiency virus gp41 polypeptide, useful in



```

CC      stop codon in AAO66275.
SQ
XX      Sequence      420 AA;
XX
XX      Query Match      99.3%; Score 290; DB 15; Length 420;
XX      Best Local Similarity 98.1%; Pred. No. 2,7e-23;
XX      Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0
XX
OY      1 NHTTWLEMDREINNTSLIHSILIESQNOQEKNEQELLELDKASLMMWPNFI 52
XX      .||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
XX      DB 155 NHTTWLEMDREINNTSLIHSILIESQNOQEKNEQELLELDKASLMMWPNFI 206
XX
XX      RESULT 5
XX      AAM00181
XX      ID AAM00181 standard; Protein; 519 AA.
XX
XX      AC AAM00181;
XX
XX      DT 12-NOV-1996 (first entry)
XX
XX      DE HIV-1 env protein.
XX
XX      KM Poliovirus; encapsidation; capsid; vaccine; genetic immunisation;
XX      KW HIV-1; human immunodeficiency virus type 1; env protein.
XX
XX      OS Human immunodeficiency virus type 1.
XX      PN WO9625173-A1.
XX      PD 22-AUG-1996.
XX
XX      PF 13-FEB-1996; 96WO-US01895.
XX
XX      PR 15-FEB-1995; 95US-0389459.
XX
XX      PA (UABR-) UAB RES FOUND.
XX
XX      PI Ansardi DC, Morrow CD, Porter DC;
XX      DR WPI; 1996-393136/39.
XX      DR N-PSDB; AAT33295.
XX
XX      PT Encapsidation of recombinant polio:virus nucleic acid for use in
XX      PT vaccines - using a polio:virus nucleic acid which lacks the PI
XX      PT capsid region and an expression system which provides the region
XX
XX      PS Disclosure: Page 61-63; 108bp; English.
XX
XX      CC cDNA sequences (AAT33293-95) respectively code for the gag, pol and
XX      CC env proteins (AAM00179-81) of HIV-1. They can be used to substitute
XX      CC the PI capsid gene of poliovirus in recombinant poliovirus nucleic
XX      CC acids (rPNAs). Such rPNAs are encapsidated by introduction into
XX      CC a host cell together with a vaccinia virus or plasmid vector
XX      CC encoding the poliovirus PI capsid precursor protein. Encapsidated
XX      CC rPNAs are useful for genetic immunisation, stimulating an immune
XX      CC response to the HIV-1 protein.
XX
XX      SQ Sequence 519 AA;
XX
OY      Query Match 99.3%; Score 290; DB 17; Length 519;
XX      Best Local Similarity 98.1%; Pred. No. 3.5e-23;
XX      Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
OY      1 NHTTWLEMDREINNTSLIHSILIESQNOQEKNEQELLELDKASLMMWPNFI 52
XX      .||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
XX      DB 420 NHTTWLEMDREINNTSLIHSILIESQNOQEKNEQELLELDKASLMMWPNFI 471
XX
XX      RESULT 6
XX      AAR69997
XX      ID AAR69997 standard; Protein; 521 AA.
XX

```

```

AC AAR69997;
XX
DT 13-SEP-1995 (first entry)
XX
DE HIV-1 env protein.
XX
KW Encapsulation; Human immunodeficiency virus; type 1; immunogenic;
KW poliovirus.
XX
OS Human immunodeficiency virus type 1.
XX
PN CA2125344-A.
XX
PD 02-JAN-1995.
XX
PE 07-JUN-1994; 94CA-2125344.
XX
PR 01-JUL-1993; 93US-0087009.
XX
PA (UABR-) UAB RES FOUND.
XX
PI Morrow CD;
XX
DR WPI; 1995-099021/14.
DR N-PSDB; AAQ80575.
XX
PT Method for encapsidating recombinant polio:virus nucleic acid -
PT useful for providing comps. to stimulate immune response.
XX
PS Disclosure; Page 38; 62pp; English.
XX
CC The sequence is that of the HIV-1 env protein. The DNA encoding
CC such protein is used in a method (claimed) to encapsulate
CC poliovirus cDNA to make it more immunogenic.
CC See also AAR6995-6.
XX
SQ Sequence 521 AA;

Query Match 99.3%; Score 290; DB 16; Length 521;
Best Local Similarity 98.1%; Pred. No. 3.5e-23;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLEESONQOEKNEOELLELDKMSLWMWNI 52
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 420 NHTTWLEMDREINNTSLIHSLEESONQOEKNEOELLELDKMSLWMWNI 471

RESULT 7
AAM43066
ID AAM43066 standard; peptide: 853 AA.
XX
AC AAM43066;
XX
DT 11-SEP-1998 (first entry)
XX
DE HIV-1 gp120 protein fragment from isolate HXB2.
XX
KW gp120 protein; purification; fractionation; ion exchange; chromatography;
KW binding affinity; CD4; hydrophobic interaction; size exclusion; vaccine.
XX
OS Human immunodeficiency virus type 1.
XX
PN US5696238-A.
XX
PD 09-DEC-1997.
XX
PE 11-MAY-1995; 95US-0439286.
XX
PR 20-AUG-1991; 91US-0684963.
PR 16-AUG-1993; 93US-0109002.
PR 09-MAY-1994; 94US-0240073.
PR 11-MAY-1995; 95US-0439286.
XX

```

```

PA (CHIR ) CHIRON CORP.
XX
PI Halgwood NL, Scandella C;
XX
DR WPI; 1998-041353/04.
XX
PT Purification of HIV gp120 - using chromatographic methods
XX
PS Disclosure; Fig 2A-W; 53pp; English.
XX
CC AAM43066-W43080 are fragments of the gp120 protein from different human
CC immunodeficiency virus type 1 (HIV-1) isolates. These proteins are used
CC in a novel method for purifying HIV gp120 so as to provide a purified
CC gp120 glycoprotein having protein/protein binding properties
CC substantially identical to natural viral HIV gp120. The method involves
CC fractionating a crude gp120 preparation containing full-length,
CC glycosylated gp120 using ion exchange chromatography so as to provide a
CC first collection of fractions. A fraction from the first collection is
CC selected that exhibits specific binding affinity for CD4 peptide,
CC thereby producing a first fractionated material. The first fractionated
CC material is fractionated by hydrophobic interaction chromatography so as
CC to provide a second collection of fractions from which a second
CC collection is selected that exhibits specific binding affinity for CD4
CC peptide. This second fraction is fractionated by size exclusion
CC chromatography so as to provide a third collection of fractions
CC exhibiting specific binding affinity for CD4 peptide, thereby providing
CC the purified gp120. The purified gp120 can be used for antibody
XX
SQ Sequence 853 AA;

Query Match 99.3%; Score 290; DB 19; Length 853;
Best Local Similarity 98.1%; Pred. No. 6.1e-23;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLEESONQOEKNEOELLELDKMSLWMWNI 52
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 621 NHTTWLEMDREINNTSLIHSLEESONQOEKNEOELLELDKMSLWMWNI 672

RESULT 8
AAP60131
ID AAP60131 standard; Protein: 856 AA.
XX
AC AAP60131;
XX
DT 26-JUN-1991 (first entry)
XX
DE Sequence of the AIDS envelope protein.
XX
KW AIDS; HIV; LAV; HTLV-III; vaccine; antibody; epitope; antigen;
KW diagnosis.
XX
OS HTLV-III.
XX
PN EP199301-A.
XX
PD 29-OCT-1986.
XX
PE 18-APR-1986; 86EP-0105371.
XX
PR 19-APR-1985; 85US-0725021.
XX
PA (HOFF ) HOFFMANN-LA ROCHE AG.
PA (USGO ) US GOVERNMENT.
PA (HEAL-) DEPT. HEALTH & HUMAN SERV.
PA (USDH ) US DEPT HEALTH & HUMAN.
XX
PI Crowl RM, Gallo RC, Reddy EP, Shaw GM, Wong-Staal FY;
XX
DR WPI; 1986-286067/44.
DR N-PSDB; AAN60128.
XX

```

PT	Envelope protein of acquired immune deficiency syndrome virus -
PT	Best Local Similarity 98.1%; Pred. No. 6.1e-23;
PT	Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX	Claim 33; Fig 6A; 46pp; English.
PS	
CC	An expression vector contg. a gene (AA60128) coding for an envelope
CC	protein of an AIDS virus, and the envelope protein of an AIDS virus
CC	(AA60131) are claimed. The vector is pref. a member of the pEV/env
CC	family, e.g. pEV1.2 or 3/env 44-640 or 205-640.
CC	
XX	
SO	Sequence 856 AA:
OY	1 NHTTWMEMREINNTSLIHSILIESONOEKNEDELLMDWASIMMNFNI 52
Db	624 NHTTWMEMREINNTSLIHSILIESONOEKNEDELLMDWASIMMNFNI 675
RESULT 9	
AAR41025	
ID	AAR41025 standard; protein; 856 AA.
XX	
AC	AAR41025;
XX	
DT	23-MAR-1994 (first entry)
XX	
DE	Selectively deglycosylated HIV-1 HXB2 env gp160 mutein C4.
XX	
KM	N-linked glycosylation; gp120; human immunodeficiency virus; type 1;
KM	envelope glycoprotein; vaccine; site-directed mutagenesis.
XX	
OS	Human immunodeficiency virus type 1 (HXB2 strain).
XX	
FH	Key Location/Qualifiers
FT	Peptide 1..30
FT	/label= signal_peptide
FT	Protein 31..511
FT	/label= gp120
FT	/note= "exterior membrane glycoprotein"
FT	512..856
FT	/label= gp41
FT	/note= "transmembrane glycoprotein"
FT	Modified-site 88
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 136
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 141
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 156
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 160
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 186
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 197
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 230
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 234
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"

FT	Modified-site	241	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	262	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	276	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	289	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	295	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	301	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	332	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	339	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	356	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	392	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	448	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	611	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	616	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	624	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	637	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	674	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	750	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	816	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Modified-site	866	/label= N-linked_glycosylation_site
FT		/note= "putative"	
FT	Misc-difference	397	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FT			
FT	Misc-difference	406	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FT			
FT	Misc-difference	463	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
FT			
XX	W09317705-A.		
XX			

PD	16-SEP-1993.	
XX		
PF	24-FEB-1993;	93WO-US01598.
XX		
PR	13-MAR-1992;	92US-0850770.
XX		
PA	(HARD )	HARVARD COLLEGE.
XX		
PI	Essex ME, Lee C, Lee T, Lee W;	
XX		
DR	WPI: 1993-303140/38.	
XX		
PT	Compsn. conf. selectively de-glycosylated HIV-1 envelope protein	
PT	- shows improved protective immune response	
XX		
PS	Claim 7; Page 15 and 23; 45pp; English.	
XX		
CC	Mutant gp160 env protein C4 is specifically claimed. The	
CC	amino acid sequence of the mutcin does not appear in the	
CC	specification; sequence RA1025 has been derived from the HIV-1 HXB2	
CC	isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHYIX,	
CC	Acc.#: P04578) and the description of C4 mutcin given in the	
CC	specification. The combination of changes made to N-linked	
CC	glycosylation sites in the C-terminal region of gp120 does not prevent	
CC	viral infectivity but the resultant selective deglycosylation enables	
CC	an immune response to be elicited by the mutcin.	
XX		
SQ	Sequence 856 AA;	
Query Match	99.3%; Score 290; DB 14; Length 856;	
Best Local Similarity	98.1%; Pred. No. 6,1e-23;	
Matches	51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	
QY	1 NHTTWLEMDREINNYTSLIHSIIIEESNOQEKNEQELLEDKWASLWMWNT 52	
DB	624 NHTTWLEMDREINNYTSLIHSIIIEESNOQEKNEQELLEDKWASLWMWNT 675	
RESULT 10		
AA	AA1026	
ID	AA1026 standard; protein; 856 AA.	
XX		
AC	AA1026;	
XX		
DT	23-MAR-1994 (first entry)	
XX		
DX	Selectively deglycosylated HIV-1 HXB2 env gp160 mutcin C5.	
DE		
XX		
KW	N-linked glycosylation; gp120; human immunodeficiency virus; type 1;	
KW	envelope glycoprotein; vaccine; site-directed mutagenesis.	
XX		
OS	Human immunodeficiency virus type 1 (HXB2 strain).	
XX		
Key	Location/Qualifiers	
FT	1..30	
FT	/label= signal_peptide	
FT	31..511	
FT	/label= gp120	
FT	/note= "exterior membrane glycoprotein"	
FT	512..856	
FT	/label= gp41	
FT	/note= "transmembrane glycoprotein"	
FT	88	
FT	/label= "N-linked.glycosylation_site	
FT	/note= "putative"	
FT	136	
FT	/label= "N-linked.glycosylation_site	
FT	/note= "putative"	
FT	141	
FT	/label= "N-linked.glycosylation_site	
FT	/note= "putative"	
FT	156	
FT	/label= "N-linked.glycosylation_site	

FT	/note= "putative"	160
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	186
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	197
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	230
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	234
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	241
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	262
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	276
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	289
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	295
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	301
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	332
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	339
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	356
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	448
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	611
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	616
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	624
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	637
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	674
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	750
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	816
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	386
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	392
FT	/label= N-1linked_glycosylation_site	
FT	/note= "putative"	392

FT eliminated by substitution of Asn by  
FT Gln at position 392"  
FT Misc-difference 397  
FT /note="putative N-linked glycosylation site  
FT eliminated by substitution of Asn by  
FT Gln at position 397"  
FT Misc-difference 406  
FT /note="putative N-linked glycosylation site  
FT eliminated by substitution of Asn by  
FT His at position 406"  
FT Misc-difference 463  
FT /note="putative N-linked glycosylation site  
FT eliminated by substitution of Asn by  
FT His at position 463"  
PN W09317705-A.  
PD 16-SEP-1993.  
XX 24-FEB-1993; 93WO-US01598.  
XX 13-MAR-1992; 92US-0850770.  
XX (HARD ) HARVARD COLLEGE.  
XX Essex ME, Lee C, Lee T, Lee W;  
XX WPI; 1993-303140/38.  
DR  
XX  
XX Compsn. conty. selectively de-glycosylated HIV-1 envelope protein  
PT - shows improved protective immune response  
XX  
XX  
PS Claim 7; page 15 and 23; 45pp; English.  
XX  
XX  
XX Mutant gp160 env protein C5 is specifically claimed. The  
CC amino acid sequence of the mutein does not appear in the  
CC specification; sequence R41026 has been derived from the HIV-1 HXB2  
CC isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIVLX,  
CC Acc. #: P04578) and the description of C5 mutein given in the  
CC specification. The combination of changes made to N-linked  
CC glycosylation sites in the C-terminal region of gp120 does not prevent  
CC viral infectivity but the resultant selective deglycosylation enables  
CC an immune response to be elicited by the mutein.  
XX  
XX  
SQ Sequence 856 AA;  
Query Match 99.3%; Score 290; DB 14; Length 856;  
Best Local Similarity 98.1%; Pred. No. 6; Ie-23;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 NHTTWEMREINNTSLIHSLEESONQOEKNEQELLEDKWASIMMNFNT 52  
DB 624 NHTTWEMREINNTSLIHSLEESONQOEKNEQELLEDKWASIMMNFNT 675  
RESULT 11  
AAR41027  
ID AAR41027 standard; protein; 856 AA.  
XX  
XX AAR41027;  
XX  
XX 23-MAR-1994 (first entry)  
XX  
XX Selectively deglycosylated HIV-1 HXB2 env gp160 mutein C6.  
XX  
XX N-linked glycosylation; gp120; human immunodeficiency virus; type 1;  
KW envelope glycoprotein; vaccine; site-directed mutagenesis.  
XX  
XX Human immunodeficiency virus type 1 (HXB2 strain).  
OS  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..30  
FT /label= signal\_peptide

FT Protein  
FT /label= gp120  
FT /note="exterior membrane glycoprotein"  
FT 512..856  
FT /label= gp41  
FT /note="transmembrane glycoprotein"  
FT 88  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 136  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 141  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 156  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 160  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 186  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 197  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 230  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 234  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 241  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 262  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 276  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 289  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 295  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 301  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 332  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 339  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 356  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 611  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 616  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 624  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 637  
FT /label= N-linked\_glycosylation\_site  
FT /note="putative"  
FT 674  
FT Modified-site

FT /label= "N-linked glycosylation\_site"  
FT /note= "putative"  
FT Modified-site 750  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT Modified-site 816  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT Misc-difference 386  
FT /note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"  
FT Misc-difference 392  
FT /note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 392"  
FT Misc-difference 397  
FT /note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"  
FT Misc-difference 406  
FT /note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"  
FT Misc-difference 448  
FT /note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 448"  
FT Misc-difference 463  
FT /note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"  
FT W09317705-A.  
FT 16-SEP-1993.  
FT 24-FEB-1993; 93MO-US01598.  
FT 13-MAR-1992; 92US-0850770.  
FT (HARD ) HARVARD COLLEGE.  
FT Essex ME, Lee C, Lee T, Lee W;  
FT WPI; 1993-303140/38.  
FT Compn. conty. selectively de-glycosylated HIV-1 envelope protein  
FT - shows improved protective immune response  
FT Claim 7; Page 15 and 23; 45pp; English.  
FT Mutant gp160 env protein C6 is specifically claimed. The  
FT amino acid sequence of the mutlein does not appear in the  
FT specification; sequence RA1027 has been derived from the HIV-1 HXB2  
FT isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIV1X,  
FT Acc.# P04578) and the description of C6 mutlein given in the  
FT specification. The combination of changes made to N-linked  
FT glycosylation sites in the C-terminal region of gp120 does not prevent  
FT viral infectivity but the resultant selective deglycosylation enables  
FT an immune response to be elicited by the mutlein.  
SQ Sequence 856 AA;  
Query Match 99.3%; Score 290; DB 14; Length 856;  
Best Local Similarity 98.1%; Pred. No. 6,1e-23;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 12  
ID AAR41028 standard; protein; 856 AA.  
XX AAR41028;  
AC AAR41028;  
DT 23-MAR-1994 (first entry)  
XX Selectively deglycosylated HIV-1 HXB2 env gp160 mutlein Q.  
DE N-linked glycosylation; gp120; human immunodeficiency virus; type 1;  
KW envelope glycoprotein; vaccine; site-directed mutagenesis.  
XX Human immunodeficiency virus type 1 (HXB2 strain).  
OS  
XX Key  
FH Location/Qualifiers  
FT 1..30  
FT Peptide  
FT /label= signal-peptide  
FT 31..511  
FT Protein  
FT /label= gp120  
FT /note= "exterior membrane glycoprotein"  
FT 512..856  
FT /label= gp41  
FT /note= "transmembrane glycoprotein"  
FT 88  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 136  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 141  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 156  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 160  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 186  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 197  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 230  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 234  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 241  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 262  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 276  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 289  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 295  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 301  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 392  
FT Modified-site  
FT /label= "N-linked-glycosylation\_site"  
FT /note= "putative"  
FT 448  
FT Modified-site

FT	/label= N-linked_glycosylation_site
ET	/note= "putative"
FR	Modified-site
FR	611
FR	/label= N-linked_glycosylation_site
FR	/note= "putative"
FR	616
FR	Modified-site
FR	/label= N-linked_glycosylation_site
FR	/note= "putative"
FR	624
FR	Modified-site
FR	/label= N-linked_glycosylation_site
FR	/note= "putative"
FR	637
FR	Modified-site
FR	/label= N-linked_glycosylation_site
FR	/note= "putative"
FR	674
FR	Modified-site
FR	/label= N-linked_glycosylation_site
FR	/note= "putative"
FR	750
FR	Modified-site
FR	/label= N-linked_glycosylation_site
FR	/note= "putative"
FR	816
FR	Modified-site
FR	/label= N-linked_glycosylation_site
FR	/note= "putative"
FR	332
FR	Misc-difference
FR	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 332"
FR	339
FR	Misc-difference
FR	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 339"
FR	356
FR	Misc-difference
FR	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 356"
FR	386
FR	Misc-difference
FR	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
FR	397
FR	Misc-difference
FR	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FR	406
FR	Misc-difference
FR	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FR	463
FR	Misc-difference
FR	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
FN	W09317705-A.
XX	16-SEP-1993.
XX	PD
XX	24-FEB-1993; 93WO-US01598.
PF	PR
XX	13-MAR-1992; 92US-0850770.
XX	(HARD ) HARVARD COLLEGE.
PA	Essex ME, Lee C, Lee T, Lee W;
PI	WP1: 1993-303140/38.
DR	Compsn. contg. selectively de-glycosylated HIV-1 envelope protein
XX	- shows improved protective immune response
PT	Claim 7; Page 18, 23 and 26; 45pp: English.
XX	Mutant gp160 env protein Q is specifically claimed. The amino acid sequence of the mutcin does not appear in the
CC	

CC	specification: sequence RA1028 has been derived from the HIV-1 HXB2
CC	isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHV1X,
CC	Acc.#: P04578) and the description of Q mtein given in the
CC	specification. The combination of changes made to N-linked
CC	glycosylation sites in the C-terminal region of gp120 does not prevent
CC	viral infectivity but the resultant selective deglycosylation enables
CC	an immune response to be elicited by the mteIn.
XX	
XX	
SQ	Sequence 856 AA:
	Query Match 99.3%; Score 290; DB 14; Length 856;
	Best Local Similarity 98.1%; Pred.No. 6,le-23;
	Matches 51; Conservative 1; MisMatches 0; Indels 0; Gaps 0.
Oy	1 NHHTMLEMREINNTSLHLSIESONOEKNEDELDPWASLMMFNI 52
	:     :     :     :     :     :     :     :     :
Db	624 NHTWMEDREINNTSLHLSIESONOEKNEDELDPWASLMMFNI 675
RESULT 13	
ID	AAR41029 standard; protein; 856 AA.
XX	
AC	AAR41029;
XX	
DT	23-MAR-1994 (first entry)
XX	
DE	Selectively deglycosylated HIV-1 HXB2 env gp160 mutein R.
XX	
N	N-linked glycosylation; gp120; human immunodeficiency virus; type 1;
KM	envelope glycoprotein; vaccine; site-directed mutagenesis.
XX	
OS	Human immunodeficiency virus type 1 (HXB2 strain).
XX	
FH	
Key	Location/Qualifiers
Peptide	1..30
/label=	"signal_peptide"
Protein	31..511
/label=	"gp120"
/note=	"exterior membrane glycoprotein"
Protein	512..856
/label=	"gp41"
/note=	"transmembrane glycoprotein"
Modified-site	88
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	136
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	141
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	156
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	160
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	186
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	197
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	230
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	234
/label=	"N-linked_glycosylation_site"
/note=	"putative"
Modified-site	241
/label=	"N-linked_glycosylation_site"
/note=	"putative"

FT	Modified-site	262	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	276	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	289	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	295	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	301	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	356	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	392	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	448	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	611	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	616	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	624	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	637	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	674	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	750	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	816	/label=	N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Misc-difference	332	/note=	"putative"	
FT					
FT	Misc-difference	339	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 332"	
FT					
FT	Misc-difference	386	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 339"	
FT					
FT	Misc-difference	386	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"	
FT					
FT	Misc-difference	397	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"	
FT					
FT	Misc-difference	406	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"	
FT					
FT	Misc-difference	463	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"	
FT					
XX					
PN					
SD					

[illegible]



```

FT Modified-site 160 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 186 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 197 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 230 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 234 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 241 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 262 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 276 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 289 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 295 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 356 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 392 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 448 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 611 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 616 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 624 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 637 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 674 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 750 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 816 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 301 /label= N-linked_glycosylation_site
FT /note= "putative"
FT Modified-site 332 /label= N-linked_glycosylation_site
FT /note= "putative N-linked glycosylation site
FT eliminated by substitution of Asn by
FT His at position 301"
FT Modified-site 339 /label= N-linked_glycosylation_site
FT /note= "putative N-linked glycosylation site
FT eliminated by substitution of Asn by
FT Gln at position 339"
FT Misc-difference 386

```

```

FT /note= "putative N-linked glycosylation site
FT eliminated by substitution of Asn by
FT Gln at position 386"
FT Misc-difference 397 /note= "putative N-linked glycosylation site
FT eliminated by substitution of Asn by
FT Gln at position 397"
FT Misc-difference 406 /note= "putative N-linked glycosylation site
FT eliminated by substitution of Asn by
FT His at position 406"
FT Misc-difference 463 /note= "putative N-linked glycosylation site
FT eliminated by substitution of Asn by
FT His at position 463"
FT W09317705-A.
FT 16-SEP-1993.
FT 24-FEB-1993; 93MO-US01598.
FT 13-MAR-1992; 92US-0850770.
FT (HARD ) HARVARD COLLEGE.
FT Essex ME, Lee C, Lee T, Lee W;
FT WPI; 1993-303140/38.
FT Compsn. contg. selectively de-glycosylated HIV-1 envelope protein
FT PT - shows improved protective immune response
FT PS Claim 7: Page 18, 23 and 26; 45pp; English.
FT CC Mutant gp160 env protein S is specifically claimed. the
FT CC amino acid sequence of the mutain does not appear in the
FT CC specification; sequence R41030 has been derived from the HIV-1 HXB2
FT CC isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIYX,
FT CC Acc. #: P04578) and the description of S mutain given in the
FT CC specification. The combination of changes made to N-linked
FT CC glycosylation sites in the C-terminal region of gp120 does not prevent
FT CC viral infectivity but the resultant selective deglycosylation enables
FT CC an immune response to be elicited by the mutain.
FT SQ Sequence 856 AA:
FT
FT Query Match 99.3%; Score 290; DB 14; Length 856;
FT Best Local Similarity 98.1%; Pred. No. 6.1e-23;
FT Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
FT
FT QY 1 NHTTWLEMDREINNTSLHSLSIESQNOOEKNEQDLLELDKWSIMNMFNI 52
FT |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
FT DB 624 NHTTWLEMDREINNTSLHSLSIESQNOOEKNEQDLLELDKWSIMNMFNI 675
FT
FT RESULT 15
FT AAR41031
FT ID AAR41031 standard; protein: 856 AA.
FT AC AAR41031;
FT XX 23-MAR-1994 (first entry)
FT DT
FT XX Selectively deglycosylated HIV-1 HXB2 env gp160 mutain T.
FT DE
FT XX N-linked glycosylation; gp120; human immunodeficiency virus; type 1;
FT KW envelope glycoprotein; vaccine; site-directed mutagenesis.
FT XX
FT OS Human immunodeficiency virus type 1 (HXB2 strain).
FT XX
FT FH Key Location/Qualifiers
FT FH Peptide 1..30

```

FT		/label= "signal_peptide
FT	Protein	31..511
FT		/label= "gp120
FT		/note= "exterior membrane glycoprotein"
FT	Protein	512..856
FT		/label= "gp41
FT		/note= "transmembrane glycoprotein"
FT	Modified-site	88
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	136
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	141
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	156
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	160
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	186
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	197
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	230
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	234
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	241
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	262
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	276
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	295
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	301
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	356
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	392
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	448
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	611
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	616
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	624
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	637
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	674
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"

FT		Modified-site	750	/Label= N-linked glycosylation-site /note= "putative"
FT				
FT		Modified-site	816	/label= N-linked glycosylation_site /note= "putative"
FT				
FT		Misc-difference	289	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 289"
FT				
FT		Misc-difference	332	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 332"
FT				
FT		Misc-difference	339	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 339"
FT				
FT		Misc-difference	386	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
FT				
FT		Misc-difference	397	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FT				
FT		Misc-difference	406	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FT				
FT		Misc-difference	463	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
FT				
XN		W09J17705-A.		
PX				
PD		16-SEP-1993.		
XX				
PF		24-FEB-1993; 93MO-USO1598.		
PR		13-MAR-1992; 92US-0850770.		
XX		(HARD ) HARVARD COLLEGE.		
PA				
PI		Essex ME, Lee C, Lee T, Lee W;		
DY		WPL: 1993-303140/38.		
DR				
XX				
XX		Compsn. contg. selectively de-glycosylated HIV-1 envelope protein - shows improved protective immune response		
PS		Claim 7; Page 18, 23 and 26; 45pp; English.		
CC				
CC		Mutant gp160 env protein T is specifically claimed. The amino acid sequence of the mutcin does not appear in the specification; sequence R4I031 has been derived from the HIV-1 HXB2 isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIVLX, Acc.# : P04578) and the description of T mutcin given in the specification. The combination of changes made to N-linked glycosylation sites in the C-terminal region of gp120 does not prevent viral infectivity but the resultant selective deglycosylation enables an immune response to be elicited by the mutcin.		
CX				
SQ		Sequence 856 AA:		
Query Match		99.3%; Score 290; DB 14; Length 856; Best Local Similarity 98.1%; Pred.No.6.le-23;		
Matches	51; Conservative	1; Mismatches	0; Indels	Gaps
OY	I NHITWLMDREINNTYTSLIHSLIEESONOEKNEDELLDKRWASIMMFNI	52		
DB	624 NHITWMEMREINNNTYSLSHSIESSONOEKNEDELLEDKRWASIMMFNI	675		

Search completed: March 18, 2003, 14:33:24  
Job time : 38.1429 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:33:36 ; Search time 12.5517 Seconds  
(without alignments)  
121.895 Million cell updates/sec

Title: US-09-877-606-3

Perfect score: 292

Sequence: 1 NHTTWLEMDREINNTSLIH.....NQELLEDKWASIMNMFNI 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Issued\_Patents\_AA:\*  
2: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	99.3	138	4	US-09-570-921-20
2	290	99.3	138	4	US-09-570-921-21
3	290	99.3	268	4	US-08-965-056-16
4	290	99.3	345	4	US-09-272-342B-7
5	290	99.3	519	1	US-08-589-446-8
6	290	99.3	519	1	US-08-444-882-8
7	290	99.3	519	2	US-08-389-459A-8
8	290	99.3	519	3	US-08-987-867A-8
9	290	99.3	856	2	US-07-916-098A-2
10	290	99.3	856	4	US-09-337-387-11
11	277	94.9	138	4	US-09-570-921-22
12	277	94.9	138	4	US-09-570-921-24
13	277	94.9	138	4	US-09-570-921-26
14	277	94.9	138	4	US-09-570-921-28
15	277	94.9	237	3	US-08-388-353-641
16	277	94.9	237	3	US-08-488-551B-641
17	277	94.9	268	4	US-08-965-056-17
18	277	94.9	268	4	US-08-965-056-18
19	277	94.9	282	5	PCT-US95-13335-1
20	277	94.9	338	3	US-08-486-099-90
21	277	94.9	338	3	US-08-360-107A-100
22	277	94.9	338	3	US-08-484-223B-90
23	277	94.9	338	3	US-08-919-597-90
24	277	94.9	338	3	US-08-475-668A-80
25	277	94.9	338	3	US-08-485-551A-90
26	277	94.9	338	3	US-08-471-913A-90
27	277	94.9	338	4	US-08-485-264A-90

28	277	94.9	338	4	US-08-474-349A-90	Sequence 90, Appl
29	277	94.9	338	4	US-08-255-208A-26	Sequence 26, Appl
30	277	94.9	345	4	US-08-817-441-49	Sequence 49, Appl
31	277	94.9	615	4	US-09-257-490-11	Sequence 11, Appl
32	277	94.9	826	1	US-08-375-510-2	Sequence 2, Appl
33	277	94.9	826	2	US-08-487-657-2	Sequence 2, Appl
34	277	94.9	839	4	US-08-472-240A-10	Sequence 10, Appl
35	277	94.9	854	4	US-09-309-572-23	Sequence 23, Appl
36	277	94.9	856	4	US-09-124-900-9	Sequence 9, Appl
37	277	94.9	861	1	US-08-127-499A-14	Sequence 14, Appl
38	277	94.9	861	1	US-08-482-847-14	Sequence 14, Appl
39	277	94.9	861	4	US-07-956-483-10	Sequence 10, Appl
40	277	94.9	861	4	US-07-956-483-16	Sequence 16, Appl
41	277	94.9	861	4	US-08-472-240A-1	Sequence 1, Appl
42	277	94.9	861	4	US-08-472-240A-7	Sequence 7, Appl
43	277	94.9	861	4	US-08-817-441-103	Sequence 103, App
44	277	94.9	863	3	US-08-463-210-11	Sequence 11, Appl
45	277	94.9	880	2	US-08-788-815-7	Sequence 7, Appl

## ALIGNMENTS

```

RESULT 1
US-09-570-921-20
; Sequence 20, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-20

Query Match          99.3%; Score 290; DB 4; Length 138;
Best Local Similarity 98.1%; Pred. No. 5.3e-26;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 NHTTWLEMDREINNTSLIHSLIEESONOEKNEQELLEDKWASIMNMFNI 52
Db      85 NHTTWLEMDREINNTSLIHSLIEESONOEKNEQELLEDKWASIMNMFNI 136

RESULT 2
US-09-570-921-21
; Sequence 21, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
; FILE REFERENCE: 106213
; CURRENT APPLICATION NUMBER: US/09/570,921
; PRIOR FILING DATE: 2000-05-15
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: FR/97/14387
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21

```

LENGTH: 138  
TYPE: PRT  
ORGANISM: Human  
US-09-570-921-21

Query Match 99.3%; Score 290; DB 4; Length 138;  
Best Local Similarity 98.1%; Pred. No. 5.3e-26;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLELDKMSLWMWFI 52  
|||||  
DB 85 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLELDKMSLWMWFI 136

RESULT 3  
US-08-965-056-16  
Sequence 16, Application US/08965056  
Patent No. 6271198  
GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted  
APPLICANT: J. Kevin Judice  
APPLICANT: Robert S. McDowell  
APPLICANT: J. Christopher Phelan  
APPLICANT: Melissa A. Starovasnik  
APPLICANT: James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
TITLE OF INVENTION: Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/965,056  
FILING DATE: 05-No. 6271198-1997  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-965-056-16

Query Match 99.3%; Score 290; DB 4; Length 268;  
Best Local Similarity 98.1%; Pred. No. 1.1e-25;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLELDKMSLWMWFI 52  
|||||  
DB 154 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLELDKMSLWMWFI 205

RESULT 4  
US-09-272-342B-7  
Sequence 7, Application US/09272342B  
Patent No. 6294341  
GENERAL INFORMATION:  
APPLICANT: YU, YEON-GYU

APPLICANT: KIM, SUNG-HOI  
APPLICANT: RYU, JAE-RYEON  
TITLE OF INVENTION: METHOD FOR DETECTING A SUBSTANCE HAVING AN ACTIVITY TO  
TITLE OF INVENTION: INHIBIT HIV INFECTION USING IMMUNOASSAY AND VARIANT  
FILE REFERENCE: 2901-0125-0  
CURRENT APPLICATION NUMBER: US/09/272,342B  
CURRENT FILING DATE: 1999-03-19  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 345  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-272-342B-7

Query Match 99.3%; Score 290; DB 4; Length 345;  
Best Local Similarity 98.1%; Pred. No. 1.5e-25;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLELDKMSLWMWFI 52  
|||||  
DB 113 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLELDKMSLWMWFI 164

RESULT 5  
US-08-589-446-8  
Sequence 8, Application US/08589446  
Patent No. 5614413  
GENERAL INFORMATION:

APPLICANT: Morrow, Casey D.  
TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC  
TITLE OF INVENTION: ACID AND METHODS OF MAKING AND  
TITLE OF INVENTION: USING SAME  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/589,446  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION NUMBER: US/08/087,009  
APPLICATION NUMBER: US/08/087,009

ATTORNEY/AGENT INFORMATION:  
NAME: Gearty III, William C.  
REGISTRATION NUMBER: 31,359  
REFERENCE/DOCKET NUMBER: UAG-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-589-446-8

Query Match 99.3%; Score 290; DB 1; Length 519;  
Best Local Similarity 98.1%; Pred. No. 2.4e-25;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-987-867A-8

Query Match 99.3%; Score 290; DB 3; Length 519;  
Best Local Similarity 98.1%; Pred. No. 2.4e-25;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWMEFI 52  
Db 420 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWMEFI 471

RESULT 9  
US-07-916-098A-2  
Sequence 2, Application US/07916098A  
Patent No. 5871732  
GENERAL INFORMATION:  
APPLICANT: BURKLY, LINDA C.  
APPLICANT: CHISHOLM, PATRICIA L.  
APPLICANT: THOMAS, DAVID W.  
APPLICANT: ROSA, MARGARET D.  
APPLICANT: ROSA, JOSEPH J.  
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN  
TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.  
STREET: 10 SOUTH WACKER DRIVE  
CITY: CHICAGO  
STATE: ILLINOIS  
COUNTRY: U.S.A.  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/916.098A  
FILING DATE: July 24, 1992  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08843  
FILING DATE: No. 5871732ember 27, 1991  
CLASSIFICATION: 424  
APPLICATION NUMBER: 07/618,542  
FILING DATE: No. 5871732ember 27, 1990  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: JOHN J. MC DONNELL  
REGISTRATION NUMBER: 26,949  
REFERENCE/DOCKET NUMBER: 92,310-G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 715-1234  
TELEFAX: (312) 715-1234  
TELEX: 910/221-5317  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 856 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-916-098A-2

Query Match 99.3%; Score 290; DB 2; Length 856;  
Best Local Similarity 98.1%; Pred. No. 4.3e-25;

Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWMEFI 52  
Db 624 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWMEFI 675

RESULT 10  
US-09-337-387-11  
Sequence 11, Application US/09337387  
Patent No. 6420545  
GENERAL INFORMATION:  
APPLICANT: HOXIE, James A.  
APPLICANT: LABRANCHE, Celia C.  
APPLICANT: DOMS, Robert W.  
APPLICANT: HOFFMAN, Trevor L.  
TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
TITLE OF INVENTION: THERAPEUTICS  
FILE REFERENCE: Hoxie 9596-104U1 (0282)  
CURRENT APPLICATION NUMBER: US/09/337,387  
CURRENT FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: US 09/317,556  
PRIOR FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 11  
LENGTH: 856  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-337-387-11

Query Match 99.3%; Score 290; DB 4; Length 856;  
Best Local Similarity 98.1%; Pred. No. 4.3e-25;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWMEFI 52  
Db 624 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWMEFI 675

RESULT 11  
US-09-570-921-22  
Sequence 22, Application US/09570921  
Patent No. 6455265  
GENERAL INFORMATION:  
APPLICANT: SERRES, PIERRE-FRANCOIS  
TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE  
TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION  
FILE REFERENCE: 106213  
CURRENT APPLICATION NUMBER: US/09/570,921  
CURRENT FILING DATE: 2000-05-15  
PRIOR APPLICATION NUMBER: PCT/FR98/02447  
PRIOR FILING DATE: 1998-11-17  
PRIOR APPLICATION NUMBER: FR/97/14387  
PRIOR FILING DATE: 1997-11-17  
NUMBER OF SEQ ID NOS: 144  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 22  
LENGTH: 138  
TYPE: PRT  
ORGANISM: Human  
US-09-570-921-22

Query Match 94.9%; Score 277; DB 4; Length 138;  
Best Local Similarity 94.2%; Pred. No. 1.6e-24;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWMEFI 52  
Db 85 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWMEFI 136

RESULT 12



```

US-09-570-921-24
: Sequence 24, Application US/09570921
: Patent No. 6455265
: GENERAL INFORMATION:
: APPLICANT: SERIES: PIERRE-FRANCOIS
: TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
: TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
: FILE REFERENCE: 106213
: CURRENT APPLICATION NUMBER: US/09/570,921
: CURRENT FILING DATE: 2000-05-15
: PRIOR APPLICATION NUMBER: PCT/FR98/02447
: PRIOR FILING DATE: 1998-11-17
: PRIOR APPLICATION NUMBER: FR/97/14387
: PRIOR FILING DATE: 1997-11-17
: NUMBER OF SEO ID NOS: 144
: SOFTWARE: PatentIn Ver. 2.1
: SEO ID NO 24
: LENGTH: 138
: TYPE: PRT
: ORGANISM: Human
: US-09-570-921-24

```

Query Match 94.9%; Score 277; DB 4; Length 138;  
 Best Local Similarity 94.2%; Pred. No. 1,6e-24;  
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0

```

RESULT 13
US-09-570-921-26
; Sequence 26, Application US/09570921
; Patent No. 6455265
GENERAL INFORMATION:
APPLICANT: SERRES, PIERRE-FRANCOIS
TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
FILE REFERENCE: 106213
CURRENT FILING DATE: 2000-05-15
PRIOR APPLICATION NUMBER: FR/FR98/02447
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: FR/97/14387
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 26
LENGTH: 138
TYPE: PRF
ORGANISM: Human
US-09-570-921-26

Query Match          94.9%; Score 277; DB 4; Length 138;
Best Local Similarity 94.2%; Pred. No. 1.6e-24;
Matches    49; Conservative   2; Mismatches     1; Indels      0; Gaps      0.

QY       1 NHTTWLEWDREINNTSLIHSILIESQNQQEKNEDELLELDKWSLMMWNT 52
           |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db        85 NNMTWMEDRDREINNTSLIHSILIESQNQQEKNEDELLELDKWSLMMWFNI 136

RESULT 14
US-09-570-921-58
; Sequence 58, Application US/09570921
; Patent No. 6455265
GENERAL INFORMATION:
APPLICANT: SERRES, PIERRE-FRANCOIS
TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
FILE REFERENCE: 106213
CURRENT APPLICATION NUMBER: US/09/570,921

```

```

? CURRENT FILING DATE: 2000-05-15
? PRIOR APPLICATION NUMBER: PCT/FR98/02447
? PRIOR FILING DATE: 1998-11-17
? PRIOR APPLICATION NUMBER: FR/97/14387
? PRIOR FILING DATE: 1997-11-17
? NUMBER OF SEO ID NOS: 144
? SOFTWARE: PatentIn Ver. 2.1
? SEO ID NO 58
? LENGTH: 138
? TYPE: PRT
? ORGANISM: Human
?
US-09-570-921-58

```

	Query Match	94.9%	Score 277;	DB 4;	length 138;
	Best Local Similarity	94.2%;	Pred. NO.	1.6e-24;	
	Matches 49;	Conservative 2;	Mismatches 1;	IndeIs 0;	Gaps 0;
Oy	1 NHTTLEMDREINNTSLHSLSIESQOQNEBDELLFDKMAASLMMNFTI	52			
	:   :				
Dd	85 NNMTTEMDRELNNTSLSHLSIESQOQNEBDELLFDKMAASLMMNFTI	136			

```

RESULT 15
US-08-388-353-641
; Sequence 641, Application US/08388353
; Patent No. 6010895
; GENERAL INFORMATION:
; APPLICANT: Deacon, Nicholas J.
; APPLICANT: Leamont, Jennifer C.
; APPLICANT: McPhee, Dale A.
; APPLICANT: Crowe, Suzanne
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
; NUMBER OF SEQUENCES: 800
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scully, Scott, Murphy & Presser
; STREET: 400 Garden City Plaza
; CITY: Garden City
; STATE: New York
; COUNTRY: United States
; ZIP: 11530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/388,353
; FILING DATE: 14-FEB-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Digiglio, Frank S.
; REGISTRATION NUMBER: 31,346
; REFERENCE/DOCKET NUMBER: 9606
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 742-4343
; TELEFAX: (516) 742-4366
; TELEX: 230 901 SANS UR
; INFORMATION FOR SEQ ID NO: 641:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 237 amino acids
; type: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-388-353-641

Query Match          94.9%; Score 277; DB 3; Length 237;
Best Local Similarity 94.2%; Pred. No. 2.9e+24;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
```

Db 5 NNMWTMEWDREINNYYTSLIHSLIEESONQOEKNEQELLELDKWSLWNPNI 56

Search completed: March 18, 2003, 14:50:04  
Job time : 13.5517 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:36:16 ; Search time 12.0394 Seconds  
(without alignments)  
199.079 Million cell updates/sec

Title: US-09-877-606-3  
Perfect score: 292  
Sequence: 1 NHTTWLEWDREINNYTSLIH.....NEQELLEDKMASLNMWENI 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5.

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published\_Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PC7\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/PC705\_PUBCOMB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	99.3	268	10	US-09-854-816-16
2	290	99.3	519	10	US-09-756-551A-8
3	290	99.3	856	10	US-09-476-242-1
4	277	94.9	56	10	US-09-779-451-4
5	277	94.9	268	10	US-09-854-816-17
6	277	94.9	268	10	US-09-854-816-18
7	277	94.9	345	10	US-09-779-451-8
8	277	94.9	853	9	US-10-003-035-33
9	277	94.9	1101	9	US-10-003-035-53
10	277	94.9	1186	9	US-10-003-035-55
11	274	93.8	268	10	US-09-854-816-19
12	260	89.0	268	10	US-09-854-816-46
13	259	88.7	268	10	US-09-854-816-13
14	256	87.7	46	10	US-09-779-451-41
15	255	87.3	269	10	US-09-854-816-43
16	254	87.0	268	10	US-09-854-816-9
17	253	86.6	269	10	US-09-854-816-12
18	252	86.3	269	10	US-09-854-816-28
19	251	86.0	268	10	US-09-854-816-26

20	251	86.0	619	10	US-09-891-609-4	Sequence 4, Appl
21	251	86.0	646	10	US-09-891-609-2	Sequence 2, Appl
22	251	86.0	847	10	US-09-476-242-2	Sequence 2, Appl
23	250	85.6	233	10	US-09-854-816-50	Sequence 50, Appl
24	248	84.9	269	10	US-09-854-816-44	Sequence 44, Appl
25	247	84.6	268	10	US-09-854-816-8	Sequence 8, Appl
26	246	84.2	268	10	US-09-854-816-41	Sequence 41, Appl
27	243	83.2	267	10	US-09-854-816-38	Sequence 38, Appl
28	243	83.2	269	10	US-09-854-816-32	Sequence 32, Appl
29	243	83.2	269	10	US-09-854-816-34	Sequence 34, Appl
30	243	83.2	269	10	US-09-854-816-45	Sequence 45, Appl
31	242	82.9	267	10	US-09-854-816-15	Sequence 6, Appl
32	242	82.9	269	10	US-09-854-816-6	Sequence 11, Appl
33	241	82.5	267	10	US-09-854-816-11	Sequence 11, Appl
34	241	82.5	269	10	US-09-854-816-42	Sequence 42, Appl
35	240	82.2	269	10	US-09-854-816-33	Sequence 33, Appl
36	238	81.5	268	10	US-09-854-816-10	Sequence 10, Appl
37	238	81.5	268	10	US-09-854-816-35	Sequence 35, Appl
38	238	81.5	268	10	US-09-854-816-72	Sequence 72, Appl
39	238	81.5	269	10	US-09-854-816-20	Sequence 20, Appl
40	238	81.5	269	10	US-09-854-816-31	Sequence 31, Appl
41	237	81.2	233	10	US-09-854-816-49	Sequence 49, Appl
42	236	80.8	145	12	US-10-000-321-11	Sequence 11, Appl
43	236	80.8	269	10	US-09-854-816-21	Sequence 21, Appl
44	236	80.8	269	10	US-09-854-816-25	Sequence 25, Appl
45	236	80.8	269	10	US-09-854-816-30	Sequence 30, Appl

#### ALIGNMENTS

RESULT 1  
US-09-854-816-16  
Sequence 16, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Pheelan  
Melissa A. Starovasnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 Inch, 1.44 MB floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854, 816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-854-816-16

Query Match 99.3%; Score 290; DB 10; Length 268;  
Best Local Similarity 98.1%; Pred. No. 7.5e-25;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLSIESQNOQKNEQELLELDKWSLWVFN 52  
Db 154 NHTTWLEMDREINNYTSLIHSLSIESQNOQKNEQELLELDKWSLWVFN 205

RESULT 2  
US-09-756-551A-8  
Sequence 8, Application US/09756551A  
Patent No. US20020051768A1  
GENERAL INFORMATION:  
APPLICANT: C. Morrow et al.  
TITLE OF INVENTION: ENCAPSATED RECOMBINANT VIRAL  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 STATE STREET  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/756,551A  
FILING DATE: 08-JAN-2001  
CLASSIFICATION:  
APPLICATION DATA:  
PRIOR APPLICATION NUMBER: 08/376,184  
FILING DATE: 17-AUG-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/987,867  
FILING DATE: 09-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/389,459  
FILING DATE: 15-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/087,009  
FILING DATE: 01-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Laito, Peter C.  
REGISTRATION NUMBER: 32,360  
REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-756-551A-8

Query Match 99.3%; Score 290; DB 10; Length 519;  
Best Local Similarity 98.1%; Pred. No. 1.6e-24;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLSIESQNOQKNEQELLELDKWSLWVFN 52  
Db 420 NHTTWLEMDREINNYTSLIHSLSIESQNOQKNEQELLELDKWSLWVFN 471

RESULT 3  
US-09-476-242-1  
Sequence 1, Application US/09476242  
Patent No. US20020146683A1  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: HARTOG, Karin  
TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
FILE REFERENCE: 1605,002  
CURRENT APPLICATION NUMBER: US/09/476,242  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 856  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-476-242-1

Query Match 99.3%; Score 290; DB 10; Length 856;  
Best Local Similarity 98.1%; Pred. No. 2.7e-24;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLSIESQNOQKNEQELLELDKWSLWVFN 52  
Db 624 NHTTWLEMDREINNYTSLIHSLSIESQNOQKNEQELLELDKWSLWVFN 675

RESULT 4  
US-09-779-451-4  
Sequence 4, Application US/09779451  
Patent No. US20020094521A1  
GENERAL INFORMATION:  
APPLICANT: Wild, Carl T.  
TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors  
FILE REFERENCE: 1900,0300003  
CURRENT APPLICATION NUMBER: US/09/779,451  
PRIOR FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: US 60/235,901  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/181,543  
PRIOR FILING DATE: 2000-02-10  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-779-451-4

Query Match 94.9%; Score 277; DB 10; Length 56;  
Best Local Similarity 94.2%; Pred. No. 3.6e-24;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLSIESQNOQKNEQELLELDKWSLWVFN 52  
Db 2 NHTTWLEMDREINNYTSLIHSLSIESQNOQKNEQELLELDKWSLWVFN 53

RESULT 5  
US-09-854-816-17  
Sequence 17, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted

J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Phd., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

US-09-854-816-17

SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-854-816-17

Query Match 94.9%; Score 277; DB 10; Length 268;  
Best Local Similarity 94.2%; Pred. No. 2e-23;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 52  
|:|||||  
DB 154 NNMWTWMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 205

RESULT 6  
US-09-854-816-18  
Sequence 18, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Phd., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

US-09-854-816-18

SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-854-816-18

Query Match 94.9%; Score 277; DB 10; Length 268;  
Best Local Similarity 94.2%; Pred. No. 2e-23;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 52  
|:|||||  
DB 154 NNMWTWMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 205

RESULT 7  
US-09-779-451-8  
Sequence 8, Application US/09779451  
Patent No. US20020094521A1  
GENERAL INFORMATION:  
APPLICANT: Wild, Carl T.  
APPLICANT: Allaway, Graham P.  
TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors  
FILE REFERENCE: 1900.0300003  
CURRENT APPLICATION NUMBER: US/09/779,451  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: US 60/235,901  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/181,543  
PRIOR FILING DATE: 2000-02-10  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 8  
LENGTH: 345  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1

US-09-779-451-8

Query Match 94.9%; Score 277; DB 10; Length 345;  
Best Local Similarity 94.2%; Pred. No. 2.7e-23;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 52  
|:|||||  
DB 113 NNMWTWMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 164

RESULT 8  
US-10-003-035-33  
Sequence 33, Application US/10003035  
Patent No. US2002015127A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Danher

```

      : PRIOR FILING DATE: 2001-06-04
      : NUMBER OF SEQ ID NOS: 75
      : SOFTWARE: PatentIn version 3.1
      : SEQ ID NO 55
      : LENGTH: 1186
      : TYPE: PRT
      : ORGANISM: Artificial sequence
      : FEATURE:
      : OTHER INFORMATION: Modified Env/Tat/Rev
US-10-003-035-55

Query Match          94.9%; Score 277; DB 9; Length 1186;
Best Local Similarity 94.2%; Pred. No. 1,le-22;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0.

OY      1 NHTTWLEMDREINNTYSLIHSLESQNOOEKNEDELLELDKWSLMMWFI 52
       1::|:|||||||
Db      986 NNMTWMEDREINNTYSLIHSLESQNOOEKNEDELLELDKWSLMMWFI 1037

RESULT 11
US-09-854-816-19
: Sequence 19, Application US/09854816
: Patent No. US20020151473A1
GENERAL INFORMATION:
APPLICANT: Andrew C. Braisted
           J. Kevin Judice
           Robert S. McDowell
           J. Christopher Phelan
           Melissa A. Starovasnik
           James A. Wells
TITLE OF INVENTION: Constraining Helical Peptides and Methods of
                    Making Same
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSER: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/965,056
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, PhD., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: PL005N2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-854-816-19

Query Match          93.8%; Score 274; DB 10; Length 268;
Best Local Similarity 92.3%; Pred. No. 4.4e-23;
Matches 48; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

1 NHTTWLEMDREINNTYSLIHSLESQNOOEKNEDELLELDKWSLMMWFI 52

```

[illegible]

```

:
: NUMBER OF SEQUENCES: 113
: Making Same
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Genentech, Inc.
: STREET: 1 DNA Way
: CITY: South San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94080
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
:
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Winpatin (Genentech)
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/854, 816
: FILING DATE: 15-May-2001
: CLASSIFICATION: <unknown>
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/965, 056
: FILING DATE: <unknown>
:
: ATTORNEY/AGENT INFORMATION:
: NAME: Torchia, Phd., Timothy E.
: REGISTRATION NUMBER: 36,700
: REFERENCE/DOCKET NUMBER: P1005R2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 650/225-8674
: TELEFAX: 650/952-9881
:
: INFORMATION FOR SEQ ID NO: 13:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 268 amino acids
: TYPE: Amino Acid
: TOPOLOGY: Linear
:
: SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-854-816-13
:
:
: Query Match 88.7%: Score 259; DB 10; Length 268;
: Best Local Similarity 86.5%: Pred. No. 2e-21;
: Matches 45; Conservative 6; Mismatches 1; Indels 0; Gaps 0.
:
:
: 1 NHTTWLEMDREINNTSLIHSLSIESQNGEENQELLELDKWSLMMNFI 52
: :||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 154 DNMTWLEMDREINNTSLIHSLSIESQNGEENQELLELDKWSLMMNFDI 205
:
: RESULT 14
: US-09-779-451-41
: Sequence 41, Application US/09779451
: Patent No. US20020094521A1
:
: GENERAL INFORMATION:
: APPLICANT: Wild, Carl T.
: TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors
: FILE REFERENCE: 1900.030003
: CURRENT APPLICATION NUMBER: US/09/779,451
: CURRENT FILING DATE: 2001-08-17
: PRIOR APPLICATION NUMBER: US 60/235,901
: PRIOR FILING DATE: 2000-09-28
: PRIOR APPLICATION NUMBER: US 60/181,543
: PRIOR FILING DATE: 2000-02-10
: NUMBER OF SEQ ID NOS: 77
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 41
: LENGTH: 46
: TYPE: PRT
: ORGANISM: Human immunodeficiency virus type 1
US-09-779-451-41
:
: Query Match 87.7%: Score 256; DB 10; Length 46;
: Best Local Similarity 97.8%: Pred. No. 6e-22;
: Matches 45; Conservative 1; Mismatches 0; Indels 0; Gaps 0.
:
:
: 5 WLEMDREINNTSLIHSLSIESQNGEENQELLELDKWSLMMNFI 50

```





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:35:13 ; Search time 2425.56 Seconds

(without alignments)  
623.917 Million cell updates/sec

Title: US-09-877-606-3

Perfect score: 292  
Sequence: 1 NHTTWLEMDREINNTSLIH.....NEQELLEDKNASLMMWNI 52

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+g2n.model -DEV=xlh  
-O=/cgn2.1/USPTO/US09877606/runat\_18032003\_134800\_17655/app.query.fasta.1.1059  
-DB=GenEmbl -QFMT=fastap -SUFFIX=trige -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45  
-LOCALIG=200 -THR.SCORE=DCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=8to -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09877606.ecgn.1.1.3852\_runat\_18032003\_134800\_17655 -NCPU=6 -ICPU=3  
-NO\_XLPPX -NO\_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl :  
1: gb\_da :  
2: gb\_hg :  
3: gb\_in :  
4: gb\_om :  
5: gb\_ov :  
6: gb\_pat :  
7: gb\_ph :  
8: gb\_pl :  
9: gb\_pr :  
10: gb\_ro :  
11: gb\_sts :  
12: gb\_sy :  
13: gb\_un :  
14: gb\_vl :  
15: em\_da :  
16: em\_fun :  
17: em\_hum :  
18: em\_in :  
19: em\_mu :  
20: em\_om :  
21: em\_of :  
22: em\_ov :  
23: em\_pat :  
24: em\_ph :  
25: em\_pl :  
26: em\_ro :  
27: em\_sts :  
28: em\_un :

29: em\_vl :  
30: em\_hg\_hum :  
31: em\_hg\_inv :  
32: em\_hg\_other :  
33: em\_hg\_mus :  
34: em\_hg\_pln :  
35: em\_hg\_rod :  
36: em\_hg\_mam :  
37: em\_hg\_vrt :  
38: em\_sy :  
39: em\_hggo\_hum :  
40: em\_hggo\_mus :  
41: em\_hggo\_other :

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	99.3	1098	9 HDWIGGZGP	M18404 Human Igg2
2	290	99.3	1260	6 AR000534	AR000534 Sequence
3	290	99.3	1568	6 AR044676	AR044676 Sequence
4	290	99.3	1568	6 I38637	I38637 Sequence 7
5	290	99.3	1568	6 I40608	I40608 Sequence 7
6	290	99.3	2571	6 AR035189	AR035189 Sequence
7	290	99.3	2571	6 AR358143	AR358143 HIV-1 clo
8	290	99.3	2571	14 AR358144	AR358144 HIV-1 clo
9	290	99.3	2573	14 AR358140	AR358140 HIV-1 clo
10	290	99.3	2573	14 AR358142	AR358142 HIV-1 clo
11	290	99.3	2680	14 HIV1012032	U12032 Human Immun
12	290	99.3	2681	14 HIV1012033	U12033 Human Immun
13	290	99.3	2691	14 HIV1012030	U12030 Human Immun
14	290	99.3	2692	14 HIV1012031	U12031 Human Immun
15	290	99.3	2694	6 I83400	I83400 Sequence 1
16	290	99.3	3084	6 I83401	I83401 Sequence 2
17	290	99.3	3156	6 E01088	E01088 Nucleic acd
18	290	99.3	3156	14 HIVHXB3	M14100 Human Immun
19	290	99.3	6031	12 XXU19867	U19867 Cloning vec
20	290	99.3	6460	6 AX427796	AX427796 Sequence
21	290	99.3	6505	6 AX427777	AX427777 Sequence
22	290	99.3	6505	6 AX427778	AX427778 Sequence
23	290	99.3	6532	6 AX427831	AX427831 Sequence
24	290	99.3	6538	6 AX427792	AX427792 Sequence
25	290	99.3	6538	6 AX427795	AX427795 Sequence
26	290	99.3	6541	6 AX427793	AX427793 Sequence
27	290	99.3	6577	6 AX427794	AX427794 Sequence
28	290	99.3	6610	6 AX427828	AX427828 Sequence
29	290	99.3	6610	6 AX427829	AX427829 Sequence
30	290	99.3	6613	6 AX427830	AX427830 Sequence
31	290	99.3	6616	6 AX427789	AX427789 Sequence
32	290	99.3	6619	6 AX427786	AX427786 Sequence
33	290	99.3	6619	6 AX427790	AX427790 Sequence
34	290	99.3	6634	6 AX427791	AX427791 Sequence
35	290	99.3	6646	6 AX427779	AX427779 Sequence
36	290	99.3	6649	6 AX427827	AX427827 Sequence
37	290	99.3	6655	6 AX427787	AX427787 Sequence
38	290	99.3	6658	6 AX427788	AX427788 Sequence
39	290	99.3	6670	6 AX427780	AX427780 Sequence
40	290	99.3	6688	6 AX427824	AX427824 Sequence
41	290	99.3	6691	6 AX427821	AX427821 Sequence
42	290	99.3	6691	6 AX427825	AX427825 Sequence
43	290	99.3	6697	6 AX427783	AX427783 Sequence
44	290	99.3	6706	6 AX427826	AX427826 Sequence
45	290	99.3	6709	6 AX427781	AX427781 Sequence

RESULT 1

## ALIGNMENTS

HUMIG2GP  
LOCUS HUMIG2GP 1098 bp DNA linear PRI 08-NOV-1994  
DEFINITION Human IgG2 lambda antibody (IIB8.env reactive) gp41 coding region  
DNA.  
ACCESSION M18404  
VERSION M18404.1 GI:184839  
KEYWORDS antibody; glycoprotein.  
SOURCE Human (HIV seropositive male) DNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE  
AUTHORS Banatour,B., Rosenthal,K., Rabin,L., Sharma,V., Young,L.,  
TITILE Fernandez,J., Engleman,E., McGrath,M., Reyes,G. and Lifson,J.  
Characterization and epitope mapping of a human monoclonal antibody  
reactive with the envelope glycoprotein of human immunodeficiency  
virus  
JOURNAL J. Immunol. 139 (12), 4027-4033 (1987)  
MEDLINE 88088754  
PUBMED 2447158  
FEATURES  
Source Location/Qualifiers  
1..1098  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="14q32.33"  
gene 1..387  
/gene="IGHG2"  
mal\_peptide 1..387  
/gene="IGHG2"  
/note="gp41; G00-119-338"  
BASE COUNT 337 a 209 c 295 g 257 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 2.3e-28 Length: 1098  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: Gaps: 0  
US-09-877-606-3 (1-52) x HUMIG2GP (1-1098)  
QY 1 AsnHsThrThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrThrSerLeuIleHis 20  
|||||  
Db 337 AATCACACGACCTGGATGGAGTGGACAGAAATTAACAATTACACAAAGCTTAATACAC 396  
QY 21 SerLeuIleGluIleGluSerGlnAsnGlnGlnIleGluIleGluIleGluIleGlu 40  
|||||  
Db 397 TCCTTAATTGAAGATGCGCAAAACACCAACAAAGATGACAAAGATTAATGGAATTA 456  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 457 GATTAATGGCGCAAGTTGTGGAATGTGTTAACAATA 492  
RESULT 2  
LOCUS AR000534 1260 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 7 from patent US 5736391.  
ACCESSION AR000534  
VERSION AR000534.1 GI:3963065  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1260)  
AUTHORS Essex,M.E., Yu,X. and Lee,T.-H.  
TITILE HIV gp41 mutants  
JOURNAL Patent: US 5736391-A 7 07-APR-1998;  
FEATURES Location/Qualifiers  
1..1260  
/organism="unknown"  
BASE COUNT 398 a 214 c 357 g 291 t

ORIGIN  
Alignment Scores:  
Pred. No.: 2.65e-28 Length: 1260  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: Gaps: 0  
US-09-877-606-3 (1-52) x AR000534 (1-1260)  
QY 1 AsnHsThrThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrThrSerLeuIleHis 20  
|||||  
Db 464 AATCACACGACCTGGATGGAGTGGACAGAAATTAACAATTACACAAAGCTTAATACAC 523  
QY 21 SerLeuIleGluIleGluSerGlnAsnGlnGlnIleGluIleGluIleGluIleGlu 40  
|||||  
Db 524 TCCTTAATTGAAGATGCGCAAAACACCAACAAAGATGACAAAGATTAATGGAATTA 583  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 584 GATTAATGGCGCAAGTTGTGGAATGTGTTAACAATA 619  
RESULT 3  
LOCUS AR044676 1568 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 7 from patent US 5817512.  
ACCESSION AR044676  
VERSION AR044676.1 GI:5966141  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Morrow,C.D., Porter,D.C. and Ansardi,D.C.  
TITILE Encapsidated recombinant viral nucleic acid and methods of making  
and using same  
JOURNAL Patent: US 5817512-A 7 06-OCT-1998;  
FEATURES Location/Qualifiers  
1..1568  
/organism="unknown"  
BASE COUNT 563 a 267 c 357 g 381 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.31e-28 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: Gaps: 0  
US-09-877-606-3 (1-52) x AR044676 (1-1568)  
QY 1 AsnHsThrThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrThrSerLeuIleHis 20  
|||||  
Db 1264 AATCACACGACCTGGATGGAGTGGACAGAAATTAACAATTACACAAAGCTTAATACAC 1323  
QY 21 SerLeuIleGluIleGluSerGlnAsnGlnGlnIleGluIleGluIleGluIleGlu 40  
|||||  
Db 1324 TCCTTAATTGAAGATGCGCAAAACACCAACAAAGATGACAAAGATTAATGGAATTA 1383  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATTAATGGCGCAAGTTGTGGAATGTGTTAACAATA 1419  
RESULT 4  
LOCUS I38637 1568 bp DNA linear PAT 13-MAY-1997  
DEFINITION Sequence 7 from patent US 5614413.  
ACCESSION I38637  
VERSION I38637.1 GI:2084691

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Morrow, C.D.  
TITLE Encapsidated recombinant poliovirus nucleic acid and methods of making and using same  
JOURNAL Patent: US 5614413-A 7 25-MAR-1997;  
FEATURES Location/Qualifiers  
source 1..1568  
/organism="unknown"  
BASE COUNT 563 a 267 c 357 g 381 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.31e-28 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
Gaps: 0

US-09-877-606-3 (1-52) x 138637 (1-1568)

QY 1 AsnH1sthrTrrpLeugIUTrPAspArgGluileAsnAsnTyrrHrSerleuileHis 20  
|||||  
Db 1264 AATCAGACAGCGCTGGATGGAGTGGACAGAGAAATTACACATTAATACAC 1323

QY 21 SerleuileGluIuSerGlnAsnGlnGlnGluLysAsnGluIuLeuLeuGluLeu 40  
|||||  
Db 1324 TCCTTAATGAGAAATCGCAAAACAGCAAGAAATGACAAATTAATGGAATTA 1383

QY 41 AspLystrPalaserleuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATTAATGGCGCAAGTTGTGGAATGGTTAAACATA 1419

RESULT 5  
LOCUS 140608 1568 bp DNA linear PAT 13-MAY-1997  
DEFINITION Sequence 7 from patent US 5622705.  
ACCESSION 140608.1 GI:2082088  
VERSION 140608.1  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Morrow, C.D.  
TITLE Encapsidated recombinant poliovirus nucleic acid and methods of making and using same  
JOURNAL Patent: US 5622705-A 7 22-APR-1997;  
FEATURES Location/Qualifiers  
source 1..1568  
/organism="unknown"  
BASE COUNT 563 a 267 c 357 g 381 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.31e-28 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
Gaps: 0

US-09-877-606-3 (1-52) x 140608 (1-1568)

QY 1 AsnH1sthrTrrpLeugIUTrPAspArgGluileAsnAsnTyrrHrSerleuileHis 20  
|||||  
Db 1264 AATCAGACAGCGCTGGATGGAGTGGACAGAGAAATTACACATTAATACAC 1323

QY 21 SerleuileGluIuSerGlnAsnGlnGlnGluLysAsnGluIuLeuLeuGluLeu 40  
|||||  
Db 1324 TCCTTAATGAGAAATCGCAAAACAGCAAGAAATGACAAATTAATGGAATTA 1383

QY 41 AspLystrPalaserleuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATTAATGGCGCAAGTTGTGGAATGGTTAAACATA 1419

RESULT 5  
LOCUS 140608 1568 bp DNA linear PAT 13-MAY-1997  
DEFINITION Sequence 7 from patent US 5622705.  
ACCESSION 140608.1 GI:2082088  
VERSION 140608.1  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Morrow, C.D.  
TITLE Encapsidated recombinant poliovirus nucleic acid and methods of making and using same  
JOURNAL Patent: US 5622705-A 7 22-APR-1997;  
FEATURES Location/Qualifiers  
source 1..1568  
/organism="unknown"  
BASE COUNT 563 a 267 c 357 g 381 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.31e-28 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
Gaps: 0

US-09-877-606-3 (1-52) x 140608 (1-1568)

Db 1324 TCCTTAATGAGAAATCGCAAAACAGCAAGAAATTAATGGAATTA 1383  
QY 41 AspLystrPalaserleuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATTAATGGCGCAAGTTGTGGAATGGTTAAACATA 1419

RESULT 6  
LOCUS AR035189 2571 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5871732.  
ACCESSION AR035189  
VERSION AR035189.1 GI:5951857  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2571)  
AUTHORS Burkiy, L.C., Chisholm, P.L., Thomas, D.W., Rosa, M.D. and Rosa, J.J.  
TITLE Anti-Cd4 antibody homologs useful in prophylaxis and treatment of AIDS, ARC and HIV infection  
JOURNAL Patent: US 5871732-A 1 16-FEB-1999;  
FEATURES Location/Qualifiers  
source 1..2571  
/organism="unknown"

BASE COUNT 892 a 439 c 616 g 624 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.48e-28 Length: 2571  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
Gaps: 0

US-09-877-606-3 (1-52) x AR035189 (1-2571)

QY 1 AsnH1sthrTrrpLeugIUTrPAspArgGluileAsnAsnTyrrHrSerleuileHis 20  
|||||  
Db 1870 AATCAGACAGCGCTGGAGTGGACAGAGAAATTACACATTAATACAC 1929

QY 21 SerleuileGluIuSerGlnAsnGlnGlnGluLysAsnGluIuLeuLeuGluLeu 40  
|||||  
Db 1930 TCCTTAATGAGAAATCGCAAAACAGCAAGAAATTAATGGAATTA 1989

QY 41 AspLystrPalaserleuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1990 GATTAATGGCGCAAGTTGTGGAATGGTTAAACATA 2025

RESULT 7  
LOCUS AF358143 2571 bp DNA linear VRL 03-MAY-2001  
DEFINITION HIV-1 clone C48.1 from France envelope glycoprotein (env) pseudogene, complete sequence.  
ACCESSION AF358143  
VERSION AF358143.1 GI:13936875  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1.  
REFERENCE 1 (bases 1 to 2571)  
AUTHORS Ataman-Onal, Y., Cheynet, V. and Verrier, B.  
TITLE Mutations and transcriptional alterations associated with the downregulation of HIV-1 envelope glycoprotein expression following acute cytopathic effects  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2571)  
AUTHORS Ataman-Onal, Y., Cheynet, V. and Verrier, B.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2001) UMR 2142 CNRS/Biomerieux, CERV1-IFR74, 21 Avenue Tony Garnier, Lyon, Rhone 69365, France

FEATURES  
source Location/Qualifiers  
1..2571  
/organism="Human immunodeficiency virus type 1"  
/db\_xref="taxon:11676"  
/clone="C48.1"  
/cell\_line="HL60"  
/country="France"  
1..309  
/gene="env"  
/note="envelope glycoprotein; contains premature stop codons"  
pseudo

BASE COUNT 891 a 438 c 619 g 623 t

ORIGIN  
Alignment Scores:  
Pred. No.: 5.48e-28 Length: 2571  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-3 (1-52) x AF358143 (1-2571)

OY 1 AsnHsThrThrTLeuGluTrpAspArgGluIleAsnAsnYrThrSerLeuIleHis 20  
|||||  
Db 1870 AATCACAGACCTGGATGGATGGACAGACAATTAACATTAACACGCTTAATACAC 1929  
|||||

OY 21 SerLeuIleGluJuserGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuLeu 40  
|||||  
Db 1930 TCCTTAATTGAAGAATGCACAAACAGACAGAAAAGATGAACAATTAATGGAATTA 1989  
|||||

OY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1990 GATAAATGGCAAGTTTGGAATTTGTTAACATA 2025  
|||||

RESULT 8  
AF358144 2571 bp DNA linear VRL 03-MAY-2001  
LOCUS  
DEFINITION HIV-1 clone C48.2 from France envelope glycoprotein (env)  
ACCESSION AF358144  
VERSION AF358144  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1.  
VIRUSES: Retroviridae; Retroviridae; Lentivirus; Primate  
Lentivirus group.  
1 (bases 1 to 2571)  
REFERENCE  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Mutations and transcriptional alterations associated with the  
downregulation of HIV-1 envelope glycoprotein expression following  
acute cytopathic effects  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2571)  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2001) UMR 2142 CNRS/Biomerieux, CERVI-IFR74, 21  
Avenue Tony Garnier, Lyon, Rhone 69365, France  
Location/Qualifiers  
1..2571  
/organism="Human immunodeficiency virus type 1"  
/db\_xref="taxon:11676"  
/clone="C48.2"  
/cell\_line="HL60"  
/country="France"  
1..309  
/gene="env"  
/note="envelope glycoprotein; contains premature stop codons"

FEATURES  
source Location/Qualifiers  
1..2571  
/organism="Human immunodeficiency virus type 1"  
/db\_xref="taxon:11676"  
/clone="C48.2"  
/cell\_line="HL60"  
/country="France"  
1..309  
/gene="env"  
/note="envelope glycoprotein; contains premature stop codons"

BASE COUNT 894 a 436 c 616 g 625 t

ORIGIN  
Alignment Scores:  
Pred. No.: 5.48e-28 Length: 2571  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-3 (1-52) x AF358144 (1-2571)

OY 1 AsnHsThrThrTLeuGluTrpAspArgGluIleAsnAsnYrThrSerLeuIleHis 20  
|||||  
Db 1870 AATCACAGACCTGGATGGATGGACAGACAATTAACATTAACACGCTTAATACAC 1929  
|||||

OY 21 SerLeuIleGluJuserGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuLeu 40  
|||||  
Db 1930 TCCTTAATTGAAGAATGCACAAACAGACAGAAAAGATGAACAATTAATGGAATTA 1989  
|||||

OY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1990 GATAAATGGCAAGTTTGGAATTTGTTAACATA 2025  
|||||

RESULT 9  
AF358140 2573 bp DNA linear VRL 03-MAY-2001  
LOCUS  
DEFINITION HIV-1 clone C8.1 from France envelope glycoprotein (env)  
ACCESSION AF358140  
VERSION AF358140  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1.  
VIRUSES: Retroviridae; Retroviridae; Lentivirus; Primate  
Lentivirus group.  
1 (bases 1 to 2573)  
REFERENCE  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Mutations and transcriptional alterations associated with the  
downregulation of HIV-1 envelope glycoprotein expression following  
acute cytopathic effects  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2573)  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2001) UMR 2142 CNRS/Biomerieux, CERVI-IFR74, 21  
Avenue Tony Garnier, Lyon, Rhone 69365, France  
Location/Qualifiers  
1..2573  
/organism="Human immunodeficiency virus type 1"  
/db\_xref="taxon:11676"  
/clone="C8.1"  
/cell\_line="HL60"  
/country="France"  
1..2570  
/gene="env"  
/note="envelope glycoprotein; contains premature stop codons"  
pseudo

BASE COUNT 894 a 442 c 614 g 623 t

ORIGIN  
Alignment Scores:  
Pred. No.: 5.49e-28 Length: 2573  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-3 (1-52) x AF358140 (1-2573)

Qy 1 AsnHsthrtrhrtrpleuglutrpsapargluileasnasyrrthrSerleuilehis 20  
Db 1870 AATCACACGACCTGGATGGAGGACAGAAATTACAAATTACACAGCTTAATACAC 1929

Qy 21 Serleuileglugluserglnasnnglncglulysasnnglncglulncglulncgluleu 40  
Db 1930 TCCTTAATGAGAAATCGCAAAACACAGAAAGAAATGAAAGAAATTAATGGAATTA 1989

Qy 41 Asplystrpalaserleutrpasnrtrpheapnle 52  
Db 1990 GATAAATGGCGCAAGTTGTGGAATTGTTTAACATA 2025

RESULT 10  
AF358142  
LOCUS AF358142 2573 bp DNA linear VRL 03-MAY-2001  
DEFINITION HIV-1 clone C8.3 from France envelope glycoprotein (env)  
ACCESSION AF358142  
VERSION AF358142  
KEYWORDS AF358142.1 GI:13936874  
SOURCE human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1  
REFERENCE 1 (bases 1 to 2573)  
AUTHORS Ataman-Onal, Y., Cheynet, V. and Verrier, B.  
TITLE Mutations and transcriptional alterations associated with the downregulation of HIV-1 envelope glycoprotein expression following acute cytopathic effects  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2573)  
AUTHORS Ataman-Onal, Y., Cheynet, V. and Verrier, B.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2001) UMR 2142 CNRS/Biomerieux, CERVI-IFR74, 21 Avenue Tony Garnier, Lyon, Rhone 69365, France  
FEATURES  
source location/Qualifiers  
1..2573  
/organism="Human immunodeficiency virus type 1"  
/strain="HXB2"  
/db\_xref="taxon:11676"  
/clone="C8.3"  
/cell\_line="HL60"  
/country="France"  
1..2570  
/gene="env"  
/note="envelope glycoprotein; contains premature stop codons"  
BASE COUNT 891 a 440 c 617 g 625 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5,49e-28 Length: 2573  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-3 (1-52) x AF358142 (1-2573)

Qy 1 AsnHsthrtrhrtrpleuglutrpsapargluileasnasyrrthrSerleuilehis 20  
Db 1870 AATCACACGACCTGGATGGAGGACAGAAATTACAAATTACACAGCTTAATACAC 1929

Qy 21 Serleuileglugluserglnasnnglncglulysasnnglncglulncglulncgluleu 40  
Db 1930 TCCTTAATGAGAAATCGCAAAACACAGAAAGAAATGAAAGAAATTAATGGAATTA 1989

Qy 41 Asplystrpalaserleutrpasnrtrpheapnle 52  
Db 1990 GATAAATGGCGCAAGTTGTGGAATTGTTTAACATA 2025

Db 1990 GATAAATGGCGCAAGTTGTGGAATTGTTTAACATA 2025

RESULT 11  
HIV1012032  
LOCUS HIV1012032 2680 bp ss-RNA linear VRL 28-SEP-1995  
DEFINITION Human immunodeficiency virus type 1 isolate LM881 from infected lab worker; envelope glycoprotein (env) gene, partial cds.  
ACCESSION U12032  
VERSION U12032  
KEYWORDS U12032.1 GI:665490  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1.  
REFERENCE 1 (bases 1 to 2680)  
AUTHORS Reitz, M., Hall, L., Robert-Guroff, M., Lautenberger, J., Hahn, B., Shaw, G., Kong, L., Weiss, S., Waters, D., Gallo, R. and Blattner, W.  
TITLE Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIIB)  
JOURNAL AIDS Res. Hum. Retroviruses 10 (9), 1143-1155 (1994)  
MEDLINE 95127297  
PUBMED 7826699  
REFERENCE 2 (bases 1 to 2680)  
AUTHORS Mulder, K. E.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-1994) Katherine E. Mulder, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
COMMENT Envelope sequence, lambda clone from SalI to BamHI.  
FEATURES  
source location/Qualifiers  
1..2680  
/organism="Human immunodeficiency virus type 1"  
/isolate="LM881"  
/db\_xref="taxon:11676"  
437..2680  
/gene="env"  
437..>2680  
/gene="env"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AA16668.1"  
/db\_xref="GI:665491"  
/translation="MRVKKYQHLRRMGWMTGLGMLICSGATEKLVTVYVYGV  
WKEATTLPCASDAKAYDTEVHNVAHACVPDPDNOEVLVAVTEFMKNDVPE  
OMHEDTSLMDQSLKRCVKTTPCLVSLKCTDILNDPTNTSSSGMTEKEIKCSN  
ISTSIKGVKRETAFTKDIPIINDTSTYTLSCSTSVYTAQCPKVSFEPIPIHC  
APAGFALIKCNKNTFNGTGPCTVSVYQCHGIRPVSTQLLNGSLAESEVYIRSN  
FTDNVKTIIYQINQSVYINCTKPNNTGKRIRIDRGGRFTVIGKIGNRQACNIS  
RAKWNNTLKOIASKLREOYNNKTIIFKOSGGDLIVTSHFNGGFFPCNSTOLF  
STWPNSTGNSNTGSDRTITLPCRIKOIIMMORVKAAPISGOTIRCSNTGILL  
TROGNNNNSEIFRGGGDMRDMNSSELYKYVYKIEPLGVAPTAKRRVYOREKA  
VGTGALPLGFLGAGSTWGAASNTLVQAROLSGIVQDNNILRAEOHLLQTLV  
WGIKQDARIILAVERYLKDOQLIGWCGSGRLCTTAVPNWASNSLSLRNHTTW  
MEWDREINNTSLSHLSESONQOEKQELLEDKVAASLMMNFNTNMLMAYKIFI  
MIVGGLVGLRIVAVLSIVNRVQGYSPLSFGTHLPFGPDRREGIEEGGERDRR  
SIRLVNVS"

BASE COUNT 965 a 451 c 623 g 641 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5,72e-28 Length: 2680  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-3 (1-52) x HIV1012032 (1-2680)

Qy 1 AsnHsthrtrhrtrpleuglutrpsapargluileasnasyrrthrSerleuilehis 20  
Db 2294 AATCACACGACCTGGATGGAGGACAGAAATTACAAATTACACAGCTTAATACAC 2353

OY	21	SerLeuIIeGluGISeRcIasngIngluLySaSgIuGlInglUdeuLeuGIueu	40
Db	2354	TCTTAAATTGAAGATCCGAAACCAACAAGAATAAGATGAACAATTAATGGAAATTA	2413
OY	41	AspLysTrpAlaSerLeuTrpAsnTrpPheasnle	52
Db	2414	GATAAATGGCGCAAGTTGTGTGAATATGGTTTAACATA	2449
RESULT	12		
HIVUI12033			
LOCUS			
DEFINITION	HIVYIU12033	2681 bp ss-RNA	linear VRL 28-SEP-1995
ACCESSION	Human immunodeficiency virus type 1 isolate LW88-2 from infected lab worker, envelope glycoprotein (env) pseudogene, partial cds.		
VERSION	U12033		
KEYWORDS	U12033.1 GI:665492		
SOURCE			
ORGANISM	Human immunodeficiency virus type 1. Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate Lentivirus group.		
REFERENCE	1 (bases 1 to 2681) Reitz,M., Hall,L., Robert-Guroff,M., Lautenberger,J., Hahn,B., Shaw,G., Konig,L., Weiss,S., Waters,D., Gallo,R. and Blattner,W.		
AUTHORS	Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIB)		
JOURNAL	AIDS Res. Hum. Retroviruses 10 (3), 1143-1155 (1994)		
MEDLINE	95127297		
PUBMED	7826699		
REFERENCE	2 (bases 1 to 2681) Mulder,K.E.		
AUTHORS	Direct Submission		
TITLE	Submitted (06-JUN-1994) Katherine E. Mulder, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA		
JOURNAL	Envelope sequence, lambda clone from SalI to BamHI.		
FEATURES			
source	location/Qualifiers 1..2681 /organism="Human immunodeficiency virus type 1" /dsname="LW88-2" /id_xref="taxon:11676" /gene="env" /gene="env" /gene="env" /pseudo /codon_start=1 /product="envelope glycoprotein"		
BASE COUNT	966 a 451 c 622 g 642 t		
ORIGIN			
Alignment Scores:			
Pred. No.:	5,72e-28	Length:	2681
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	99.32%	Indels:	0
DB:	14	Gaps:	0
US-09-877-606-3 (1-52) x HIVYIU12033 (1-2681)			
OY	1	AsnHisThrThrTPLeuGLUTrPASPARGLUlleAsnAsnTYrThSerLeulleHis	20
Db	2295	AATCACAGACCTGGATGGAGTGCGACAGAGAAATTAACAATTAACAACAAGCTTAATACAC	2354
OY	21	SerLeuIIeGluGISeRcIasngIngluLySaSgIuGlInglUdeuLeuGIueu	40
Db	2355	TCTTAAATTGAAGATCCGAAACCAACAAGAATAAGATGAACAATTAATGGAAATTA	2414
OY	41	AspLysTrpAlaSerLeuTrpAsnTrpPheasnle	52
Db	2415	GATAAATGGCGCAAGTTGTGTGAATATGGTTTAACATA	2450
RESULT	13		

[illegible]

Qy 41 AsplystrPalasertleutPaspntrPheasnlle 52  
Db 2425 GATAAATGGCAAGTTGTGGAATTGGTTTACATA 2460

RESULT 14  
LOCUS HIVU12031

DEFINITION HIVU12031 2692 bp ss-RNA linear VRL 28-SEP-1995  
Human immunodeficiency virus type 1 isolate LM85 from infected lab  
worker, envelope glycoprotein (env) gene, partial cds.

ACCESSION U12031.1 GI:665488

VERSION 012031.1

KEYWORDS Human immunodeficiency virus type 1.  
Viruses; Retroviral viruses; Retroviridae; Lentivirus; Primate

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

2 (bases 1 to 2692)

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

SOURCE

gene

CDS

BASE COUNT 970 a 457 c 624 g 641 t  
ORIGIN

Alignment Scores:

Pred. No.: 5,75e-28 Length: 2692  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-3 (1-52) x HIVU12031 (1-2692)

Qy 1 AsnHISThrTrtPleuglUtrPaspArgGluIleasnAsnYrThrSerLeuIleHis 20

Db 2306 AATCACACGACCTGATGAGTGGACAGAGAAATTAACTACACAGCTTAATACAC 2365  
Qy 21 SerleuilegluglUserGlnAsnGlnGlnGlnulysAsnGlnGlnulUleuLeuGlnLeu 40  
Db 2366 TCCCTTAATTGAAGATGCAAAACCAAGAAAGATGAACAAAGATTAATTGGAATT 2425

Qy 41 AsplystrPalasertleutPaspntrPheasnlle 52

Db 2426 GATAAATGGCAAGTTGTGGAATTGGTTTACATA 2461

RESULT 15

LOCUS 183400

DEFINITION Sequence 1 from patent US 5714316.

ACCESSION 183400

VERSION 183400.1 GI:3406930

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

BASE COUNT 889 a 515 c 657 g 633 t

ORIGIN

Alignment Scores:

Pred. No.: 5,75e-28 Length: 2694  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 6 Gaps: 0

US-09-877-606-3 (1-52) x 183400 (1-2694)

Qy 1 AsnHISThrTrtPleuglUtrPaspArgGluIleasnAsnYrThrSerLeuIleHis 20

Db 1978 AATCACACGACCTGATGAGTGGACAGAGAAATTAACTACACAGCTTAATACAC 2037

Qy 21 SerleuilegluglUserGlnAsnGlnGlnGlnulysAsnGlnGlnulUleuLeuGlnLeu 40

Db 2038 TCCCTTAATTGAAGATGCAAAACCAAGAAAGATGAACAAAGATTAATTGGAATT 2097

Qy 41 AsplystrPalasertleutPaspntrPheasnlle 52

Db 2098 GATAAATGGCAAGTTGTGGAATTGGTTTACATA 2133

Search completed: March 25, 2003, 15:49:33  
Job time: 2429.56 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 25, 2003, 11:26:42 ; Search time 202.877 Seconds  
(without alignments)  
577.217 Million cell updates/sec

Title: US-09-877-606-3  
Perfect score: 292  
Sequence: 1 NHTTLEMDREINNYTSLIH.....NEQELLELDKMASLWMMFNI 52

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n model -DEV=ylh  
-Q=/cgn2.1/USPTO/spool/US09877606/runat\_18032003\_134800\_17645/app\_query.fasta\_1.1059  
-DB=N\_Geneseq\_101002 -OFMT=fastap -SUFTX=ing -MINMATCH=0.1 -LOOPLC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09877606\_6ccn\_1.1.564\_etrnat\_18032003\_134800\_17645 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_101002: \*  
1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT: \*  
2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT: \*  
3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT: \*  
4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT: \*  
5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT: \*  
6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT: \*  
7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT: \*  
8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT: \*  
9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT: \*  
10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT: \*  
11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT: \*  
12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT: \*  
13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT: \*  
14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT: \*  
15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT: \*  
16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT: \*  
17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT: \*  
18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT: \*  
19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT: \*  
20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT: \*  
21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT: \*  
22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT: \*  
23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*  
24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	99.3	1260	15 AA066275	Sequence of HIV-1
2	290	99.3	1568	16 AA080575	HIV-1 env gene. H
3	290	99.3	1568	17 AA733295	HIV-1 env gene. H
4	290	99.3	2684	19 AA093364	GMCSF/HIV-1 chimera
5	290	99.3	3084	19 AA093365	HIV-1/CD4 chimera
6	290	99.3	3156	7 AAN60128	Sequence of the en
7	290	99.3	3246	21 AAA48622	Lentivirus vector
8	290	99.3	6460	24 ABR91482	Modified HIV prote
9	290	99.3	6505	24 ABR91463	Modified HIV prote
10	290	99.3	6505	24 ABR91464	Modified HIV prote
11	290	99.3	6532	24 ABR91517	Modified HIV prote
12	290	99.3	6538	24 ABR91478	Modified HIV prote
13	290	99.3	6538	24 ABR91481	Modified HIV prote
14	290	99.3	6541	24 ABR91479	Modified HIV prote
15	290	99.3	6577	24 ABR91480	Modified HIV prote
16	290	99.3	6610	24 ABR91514	Modified HIV prote
17	290	99.3	6610	24 ABR91515	Modified HIV prote
18	290	99.3	6613	24 ABR91516	Modified HIV prote
19	290	99.3	6616	24 ABR91475	Modified HIV prote
20	290	99.3	6619	24 ABR91472	Modified HIV prote
21	290	99.3	6619	24 ABR91476	Modified HIV prote
22	290	99.3	6634	24 ABR91477	Modified HIV prote
23	290	99.3	6646	24 ABR91465	Modified HIV prote
24	290	99.3	6649	24 ABR91513	Modified HIV prote
25	290	99.3	6655	24 ABR91473	Modified HIV prote
26	290	99.3	6658	24 ABR91474	Modified HIV prote
27	290	99.3	6670	24 ABR91466	Modified HIV prote
28	290	99.3	6688	24 ABR91510	Modified HIV prote
29	290	99.3	6691	24 ABR91507	Modified HIV prote
30	290	99.3	6691	24 ABR91511	Modified HIV prote
31	290	99.3	6697	24 ABR91469	Modified HIV prote
32	290	99.3	6706	24 ABR91467	Modified HIV prote
33	290	99.3	6709	24 ABR91462	Modified HIV prote
34	290	99.3	6712	24 ABR91470	Modified HIV prote
35	290	99.3	6715	24 ABR91471	Modified HIV prote
36	290	99.3	6727	24 ABR91508	Modified HIV prote
37	290	99.3	6730	24 ABR91509	Modified HIV prote
38	290	99.3	6736	24 ABR91468	Modified HIV prote
39	290	99.3	6769	24 ABR91504	Modified HIV prote
40	290	99.3	6774	24 ABR91536	Modified HIV prote
41	290	99.3	6784	24 ABR91505	Modified HIV prote
42	290	99.3	6787	24 ABR91455	Modified HIV prote
43	290	99.3	6787	24 ABR91506	Modified HIV prote
44	290	99.3	6808	24 ABR91503	Modified HIV prote
45	290	99.3	6846	24 ABR91537	Modified HIV prote

ALIGNMENTS

RESULT 1  
AA066275  
AA066275 standard; DNA; 1260 BP.

AC AA066275;  
DT 28-DEC-1994 (first entry)

DE Sequence of HIV-1 encoding the gp41 polypeptide.

KW Human immunodeficiency virus; HIV-1; AIDS; lentivirus; TM protein;  
extracellular protein; transmembrane protein; gp41; ds.

OS HIV-1.

XX key Location/Qualifiers  
FH 3..1260  
FT CDS /\*tag= a

```
FT      /note="AAR53783"
FT      CDS
FT      1..1260
FT      /tag- b
FT      /note="AAR53784"
FT      2..1269
FT      /tag- c
FT      /product= gp41
FT      /note="encoded by bps 128..1160 (AAR53785)"
XX
XX      MO9412533-A.
XX
XX      09-JUN-1994.
XX
XX      12-JAN-1993; 93MO-US00212.
XX
XX      23-NOV-1992; 92US-0979975.
XX
XX      (HARD ) HARVARD COLLEGE.
XX
XX      Essex ME, Lee TH, Yu X;
XX
XX      WPI: 1994-200197/24.
XX      P-PSDB; AAR53783, AAR53784, AAR53785.
XX
XX      Method for treating HIV patients - comprises administration of
XX      mutated gp41 polypeptide
XX
XX      Claim 20: Fig 1: 54pp; English.
XX
XX      The inventors claim a method of treating a patient infected with HIV
XX      by administering a mutated gp41 polypeptide or a therapeutic
XX      composition comprising nucleic acid encoding the mutant gp41
XX      polypeptide in an expressible genetic construction. The mutant gp41
XX      polypeptide contains a deletion of at least one AA in at least one
XX      of the following regions of wild type gp41 (AAR53783): AAs 844-856;
XX      814-856; 796-856; 776-856; 753-856; or 710-856, effective to either
XX      disrupt viral replication or HIV or disrupt the assembly of viral
XX      Env proteins in an HIV infected cell. AAQ66275 corresp. to bps 7631-
XX      8890 of wt HIV-1.
XX
XX      Sequence 1260 BP; 398 A; 214 C; 356 G; 292 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 2.69e-29 Length: 1260
XX      Score: 290.00 Matches: 51
XX      Percent Similarity: 100.00% Conservative: 1
XX      Best Local Similarity: 98.08% Mismatches: 0
XX      Query Match: 99.32% Indels: 0
XX      DB: 15 Gaps: 0
XX
XX      US-09-877-606-3 (1-52) x AAQ66275 (1-1260)
XX
XX      1 AsnHstHrTPrLeuGluTPrAspArgLulIleAsnAsnYrThrSerLeuIleHis 20
XX      |||||||.....:|||||
XX      Db 464 AATACACGACCTGAGTGGAGTGGACAGCAAAATTACAAATTACACAGCTTAATACAC 523
XX
XX      21 SerLeuIleGluGluSerGlnAsnGlnGlnLulysAsnGlnGlnGluLeuLeuGluLeu 40
XX      |||||||.....:|||||
XX      Db 524 TCCTTAATTGAAGAATGCCAAACACAGCAAGAAAAGATGAACAAGATTTTGGAAATTA 583
XX
XX      41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
XX      |||||||.....:|||||
XX      Db 584 GATTAATGCGCAAGTTGTGGAATTTGTTAAACATA 619
XX
XX      RESULT 2
XX      ID AAQ80575 standard; DNA; 1568 BP.
XX
XX      AAQ80575;
XX
XX      13-SEP-1995 (first entry)
XX
XX      HIV-1 env gene.
XX
```

```
XX      Encapsulation: Human immunodeficiency virus; type 1; immunogenic;
XX      poliovirus; ss.
XX
XX      Human immunodeficiency virus type 1.
XX
XX      Key Location/Qualifiers
XX      CDS 7..1565
XX      /tag- a
XX
XX      CA2125344-A.
XX
XX      02-JAN-1995.
XX
XX      07-JUN-1994; 94CA-2125344.
XX
XX      01-JUL-1993; 93US-0087009.
XX
XX      (UABR-) UAB RES FOUND.
XX
XX      Tomorrow CD;
XX
XX      WPI: 1995-099021/14.
XX      P-PSDB; AAR6997.
XX
XX      Method for encapsidating recombinant polio:virus nucleic acid -
XX      useful for providing compsns. to stimulate immune response.
XX
XX      Disclosure; Page 38; 62pp; English.
XX
XX      The sequence is that of the HIV-1 env gene. The gene is used in a
XX      method (claimed) to encapsulate poliovirus cDNA to make it more
XX      immunogenic. The method involves providing a poliovirus DNA
XX      sequence which lacks the fragment encoding the protein necessary for
XX      viral encapsulation and an expression vector lacking an infectious
XX      poliovirus genome but encoding a portion of the encapsulating
XX      protein. The poliovirus DNA also encodes a HIV-1 gene, in this case
XX      the env gene, so the recombinant virus can induce a response to both
XX      poliovirus and HIV.
XX      See also AAQ80573-7.
XX
XX      Sequence 1568 BP; 563 A; 267 C; 357 G; 381 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 3.56e-29 Length: 1568
XX      Score: 290.00 Matches: 51
XX      Percent Similarity: 100.00% Conservative: 1
XX      Best Local Similarity: 98.08% Mismatches: 0
XX      Query Match: 99.32% Indels: 0
XX      DB: 16 Gaps: 0
XX
XX      US-09-877-606-3 (1-52) x AAQ80575 (1-1568)
XX
XX      1 AsnHstHrTPrLeuGluTPrAspArgLulIleAsnAsnYrThrSerLeuIleHis 20
XX      |||||||.....:|||||
XX      Db 1264 AATACACGACCTGAGTGGAGTGGACAGCAAAATTACAAATTACACAGCTTAATACAC 1323
XX
XX      21 SerLeuIleGluGluSerGlnAsnGlnGlnLulysAsnGlnGlnGluLeuLeuGluLeu 40
XX      |||||||.....:|||||
XX      Db 1324 TCCTTAATTGAAGAATGCCAAACACAGCAAGAAAAGATGAACAAGATTTTGGAAATTA 1383
XX
XX      41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
XX      |||||||.....:|||||
XX      Db 1384 GATTAATGCGCAAGTTGTGGAATTTGTTAAACATA 1419
XX
XX      RESULT 3
XX      ID AAT33295 standard; cDNA; 1568 BP.
XX
XX      AAT33295;
XX
XX      12-NOV-1996 (first entry)
XX
```

```

DE HIV-1 env gene.
XX
XX Poliovirus: encapsidation: capsid; vaccine: genetic immunisation;
KM HIV-1; human immunodeficiency virus type 1; env gene; ss.
XX
OS Human immunodeficiency virus type 1.
XX
XX Key Location/Qualifiers
FH CDS 7..1563
FT /*tag= a
XX
XX WO9625173-A1.
XX
XX PD 22-AUG-1996.
XX
XX PF 13-FEB-1996; 96MO-US01895.
XX
XX PR 15-FEB-1995; 95US-0389459.
XX
XX PA (UABR-) UAB RES FOUND.
XX
XX PI Ansardi DC, Morrow CD, Porter DC;
XX
XX DR WPI; 1996-393136/39.
XX
XX DR P-PSDB; AAM00181.
XX
XX PS Encapsidation of recombinant polio:virus nucleic acid for use in
XX PT vaccines - using a polio:virus nucleic acid which lacks the p1
XX PT capsid region and an expression system which provides the region
XX PS
XX PS Disclosure; Page 61-63; 108pp; English.
XX
XX CC CDNA sequences (AAMT33293-95) respectively code for the gag, pol and
XX CC env proteins (AAM00179-81) of HIV-1. They can be used to substitute
XX CC the p1 capsid gene of poliovirus. In recombinant poliovirus nucleic
XX CC acids (rPNAs). Such rPNAs are encapsidated by introduction into
XX CC a host cell together with a vaccinia virus or plasmid vector.
XX CC encoding the poliovirus p1 capsid precursor protein. Encapsidated
XX CC rPNAs are useful for genetic immunisation, stimulating an immune
XX CC response to the HIV-1 protein.
XX
XX SQ Sequence 1568 BP; 563 A; 267 C; 357 G; 381 T; 0 other;

Alignment Scores:
Pred. No.: 3,56e-29 Length: 1568
Score: 290.00 Matches: 51
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 0
Query Match: 99.32% Indels: 0
DB: 17 Gaps: 0
XX
XX
US-09-877-606-3 (1-52) x AAT33295 (1-1568)

QY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrhrSerLeuIleHis 20
DB 1264 AATCACAACGACCTGGATGGGTGGACAGAGAAATTAAACAACCTTAATATCAC 1223
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuGluLeu 40
DB 1324 TCCTTAATTGAAGAAATGCGCAAAACGACAGAAAGAAAGATGAACAAGATTATTGGAATT 1283
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
DB 1384 GATTAATGGCGACAGTTGTGGAATTGGTTTAACATA 1419

RESULT 4
AAV09364 standard; DNA; 2694 BP.
XX
XX AAV09364;
AC
XX
XX 13-MAY-1998 (first entry)
XX

```

```

DE GMCsF/HIV-1 chimeric nucleotide sequence.
XX
XX KM Human immunodeficiency virus; HIV; CD4; GMCsF; chimeric; delivery;
XX KM granulocyte-macrophage colony stimulating factor; env protein; ss.
XX
XX OS Chimeric - Human immunodeficiency virus type 1.
XX
XX OS Chimeric - Homo sapiens.
XX
XX PN US5714316-A.
XX
XX PD 03-FEB-1998.
XX
XX PF 04-NOV-1993; 93US-0147890.
XX
XX PR 04-NOV-1993; 93US-0147890.
XX
XX PR 21-JUN-1991; 91US-0718537.
XX
XX PA (UPE-) UNIV PENNSYLVANIA.
XX
XX PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
XX PI Levy DN, Weiner D, Williams W;
XX
XX DR WPI; 1998-129852/12.
XX
XX PT Chimeric constructs for HIV proteins - having HIV binding site to
XX PT CD4 replaced by ligand, useful for delivery of therapeutic or
XX PT diagnostic agents
XX
XX PS Claim 4; Fig 5A-B; 29pp; English.
XX
XX CC This is a chimeric GMCsF/HIV-1 nucleotide construct which comprises an
XX CC HIV gp160 envelope gene from which the sequence encoding amino acids
XX CC 392-446 or a fragment corresponding to the HIV binding site to CD4 has
XX CC been deleted and into which has been inserted a heterologous sequence
XX CC encoding a ligand selected from granulocyte-macrophage colony stimulating
XX CC factor (GMCsF). A recombinant viral particle can be produced by a host
XX CC cell transformed with a vector comprising this chimeric construct and a
XX CC regulatory sequence which directs expression of the chimeric construct.
XX CC The viral particle is incapable of targeting its native ligand receptor
XX CC as the chimeric protein lacks a complete env protein due to the deletion
XX CC of the native retroviral binding region. The recombinant viral particle
XX CC is non-pathogenic and incapable of recombination in vivo into a
XX CC pathogenic form because it lacks the portion of the env protein
XX CC responsible for targeting native retroviral cellular receptors and lacks
XX CC viral nucleic acids. The viral particles can be used for the delivery of
XX CC biological mediators such as therapeutic agents, gene products,
XX CC diagnostic labels or toxic agents. They can be used for delivery to, e.g.
XX CC pluripotent stem cells (using GMCsF ligand) or the gp120 protein of HIV
XX CC envelope (CD4 ligand).
XX
XX SQ Sequence 2694 BP; 889 A; 515 C; 657 G; 633 T; 0 other;

Alignment Scores:
Pred. No.: 7.11e-29 Length: 2694
Score: 290.00 Matches: 51
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 0
Query Match: 99.32% Indels: 0
DB: 19 Gaps: 0
XX
XX
US-09-877-606-3 (1-52) x AAV09364 (1-2694)

QY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrhrSerLeuIleHis 20
DB 1978 AATCACAACGACCTGGATGGGTGGACAGAGAAATTAAACAATTAACAAGCTTAATCAC 2037
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuGluLeu 40
DB 2038 TCCTTAATTGAAGAAATGCGCAAAACGACAGAAAGAAAGATGAACAAGATTATTGGAATT 2097
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
DB 2098 GATTAATGGCGACAGTTGTGGAATTGGTTTAACATA 2133

```

RESULT 5  
AAV09365  
ID AAV09365 standard; DNA; 3084 BP.  
XX  
AC AAV09365;  
XX  
DT 13-MAY-1998 (first entry)  
XX  
DE HIV-1/CD4 chimeric nucleotide sequence.  
XX  
KW Human immunodeficiency virus; HIV; CD4; GMCSF; chimeric; delivery;  
KM granulocyte-macrophage colony stimulating factor; env protein; ss.  
XX  
OS Chimeric - Human immunodeficiency virus type 1.  
OS Chimeric - Homo sapiens.  
XX  
PN US5714316-A.  
XX  
PD 03-FEB-1998.  
XX  
PF 04-NOV-1993; 93US-0147890.  
XX  
PR 04-NOV-1993; 93US-0147890.  
PR 21-JUN-1991; 91US-0718537.  
XX  
PA (TYPE-) UNIV PENNSYLVANIA.  
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
XX  
PI Levy DN, Weiner D, Williams W;  
XX  
DR WPI: 1998-129852/12.  
XX  
PT Chimeric constructs for HIV proteins - having HIV binding site to  
PT CD4 replaced by ligand, useful for delivery of therapeutic or  
PT diagnostic agents  
XX  
PS Claim 5; Fig 6A-C; 29pp; English.  
XX  
CC This is a chimeric HIV-1/CD4 nucleotide construct which comprises an  
CC HIV gp160 envelope gene from which the sequence encoding amino acids  
CC 392-446 or a fragment corresponding to the HIV binding site to CD4 has  
CC been deleted and into which has been inserted a heterologous sequence  
CC encoding a ligand selected from CD4. The ligand can also be selected  
CC from the granulocyte-macrophage colony stimulating factor (GMCSF). A  
CC recombinant viral particle can be produced by a host cell transformed  
CC with a vector comprising this chimeric construct and a regulatory  
CC sequence which directs expression of the chimeric construct. The viral  
CC particle is incapable of targeting its native ligand receptor as the  
CC chimeric protein lacks a complete env protein due to the deletion of the  
CC native retroviral binding region. The recombinant viral particle is  
CC non-pathogenic and incapable of recombination in vivo into a pathogenic  
CC form because it lacks the portion of the env protein responsible for  
CC targeting native retroviral cellular receptors and lacks viral nucleic  
CC acids. The viral particles can be used for the delivery of biological  
CC mediators such as therapeutic agents, gene products, diagnostic labels  
CC or toxic agents. They can be used for delivery to, e.g. pluripotent stem  
CC cells (using GMCSF ligand) or the gp120 protein of HIV envelope (CD4  
CC ligand).  
XX  
SQ Sequence 3084 BP; 1058 A; 559 C; 748 G; 719 T; 0 other;

Alignment Scores:  
Pred. No.: 8,45e-29 Length: 3084  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 19 Gaps: 0

US-09-877-606-3 (1-52) x AAV09365 (1-3084)

OY 1 AsnIstHrTrrTrrPleuGlUtrPAsPargClUleAsnAsnYrrHnSerLeuIleHis 20

DB 2368 AATCAGACAGACCTGGATGGAGTGGAGAGAGAAATTAACAACTTAATAC 2427  
OY 21 SerLeuIleGlUglSerGlnAsnGlnGlnGlnUlySAsnGlnGlnLeuLeu 40  
DB 2428 TCCTTAATTGAAGATTCGCAAAACACGACAGAAAGATGACAGATTATTGGAATTA 2487  
OY 41 AsPylStrPaLaserLeuTrPAsnTrPheAsnIle 52  
DB 2488 GATAAATGGCAGATTGTGTGGAATTGTTAAACATA 2523

RESULT 6  
AAN60128  
ID AAN60128 standard; DNA; 3156 BP.  
XX  
AC AAN60128;  
XX  
DT 26-JUN-1991 (first entry)  
XX  
DE Sequence of the envelope gene of the HTLV-III proviral genome  
DE (HXB-3).  
XX  
KW AIDS; HIV; LAV; HTLV-III; vaccine; antibody; epitope; antigen;  
KM diagnosis; ss.  
XX  
OS HTLV-III.  
XX  
FH Key Location/Qualifiers  
FH CDS 478..3048  
FT /\*tag= a  
XX  
PN EPI99301-A.  
XX  
PD 29-OCT-1986.  
XX  
PE 18-APR-1986; 86EP-0105371.  
XX  
PR 19-APR-1985; 85US-0725021.  
XX  
PA (HOFF) HOFFMANN-LA ROCHE AG.  
PA (USGO) US GOVERNMENT.  
PA (HEAL-) DEPT HEALTH & HUMAN SERV.  
PA (USDH) US DEPT HEALTH & HUMAN.  
XX  
PI Crowl RM, Gallo RC, Reddy EP, Shaw GM, Wong-Staal FY;  
XX  
DR WPI: 1986-286067/44.  
DR P-PSDB; AAP60131.  
XX  
PT Envelope protein of acquired immune deficiency syndrome virus -  
PT useful for improved testing of human blood for antibodies against  
PT virus and as antigen for vaccines  
XX  
PS Claim 35; Fig 1; 46pp; English.  
XX  
CC An expression vector contg. a gene (AAN60128) coding for an envelope  
CC protein of an AIDS virus, and the envelope protein of an AIDS virus  
CC (AAP60131) are claimed. The vector is pref. a member of the pEV/env  
CC family, e.g. pEV1.2 or 3/env 44-640 or 205-640.  
XX  
SQ Sequence 3156 BP; 1095 A; 536 C; 765 G; 760 T; 0 other;

Alignment Scores:  
Pred. No.: 8,71e-29 Length: 3156  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 7 Gaps: 0

US-09-877-606-3 (1-52) x AAN60128 (1-3156)

OY 1 AsnIstHrTrrTrrPleuGlUtrPAsPargClUleAsnAsnYrrHnSerLeuIleHis 20



Query Match:	99.32%	Indels:	0
DB:	24	Gaps:	0
US-09-877-606-3 (1-52) x ABR91482 (1-6460)			
QY 1 AsnHisThrThrPrlLeuGIuTrpAspArgGluIleAsnAsnTrpTrpSerLeuIleHis 20			
DB 3261 AACCAACACCACTGGATGGAGTGGGACCGGAGATCAACACATACCAAGCTCATTCAC 3320			
QY 21 SerLeuIleGluIuSerGlnAsnGlnGlnIuLysAsnGlnGlnIuLeuLeuGluIuLeu 40			
DB 3321 AGCGTCATGAGAGAGAGCCAGAACACAGACAGAGAGAGAACAGACGAGAGCTGAGAGCTG 3380			
QY 41 ASPLySTrPALaSerLeuTrpAsnTrpPheAsnIle 52			
DB 3381 GACAACTGGGCGACCTGTGGAAGTGTCAACATC 3416			
RESULT 9			
ABR91463			
ID ABR91463 standard; DNA; 6505 BP.			
XX ABR91463;			
AC			
XX			
DT 14-AUG-2002 (first entry)			
XX			
DE Modified HIV protein-encoding plasmid DNA #15.			
XX			
DE HIV: human immunodeficiency virus; gene; ds: circular; cyclic; anti-HIV;			
KM Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;			
KM acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.			
XX			
OS Synthetic.			
XX			
PN WO200232943-A2.			
XX			
PD 25-APR-2002.			
XX			
PF 14-AUG-2001; 2001WO-US25721.			
XX			
PR 14-AUG-2000; 2000US-225097P.			
PR 14-NOV-2000; 2000US-252115P.			
PR 28-MAR-2001; 2001US-279257P.			
XX			
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.			
PA (CHAD/) CHADRAABARTI B K.			
XX			
PI Nabel GJ, Huang Y;			
XX			
DR WPI: 2002-452382/48.			
XX			
PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and			
PT its encoded protein, useful as vaccines for genetic or protein			
PT immunisation for acquired immunodeficiency syndrome or human			
PT immunodeficiency virus infection -			
XX			
PS Disclosure; Page 409-411; 794pp; English.			
XX			
XX The invention relates to a nucleic acid molecule encoding a modified HIV			
CC (human immunodeficiency virus) protein. The nucleic acid molecule and its			
CC genetic or protein immunisation to a host, respectively. In particular			
CC these are useful for ameliorating the symptoms of acquired			
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an			
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,			
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent			
CC compositions are useful for treating or preventing HIV infections or			
CC AIDS. Sequences ABR91449-ABR91624 represent modified HIV protein-encoding			
CC plasmid DNA sequences of the invention.			
XX			
SO Sequence 6505 BP; 1638 A; 1775 C; 1638 G; 1454 T; 0 other:			
Alignment Scores:			
Pred. No.:	2.2e-28	Length:	6505

	Score:	290.00	Matches:	51
	Percent Similarity:	100.00%	Conservative:	1
	Pest Local Similarity:	98.08%	Mismatch:	0
	Query Match:	99.32%	Indels:	0
DB:		24	Gaps:	0
US-09-877-606-3 (1-52) x ABR91463 (1-6505)				
Qy	1 AsnHsrTrhrrTTPLeuglurTPAsPaRgGluIlEasnaSntYrThSerleulleHis 20 			
Dd	3319 AACCAACACCACCGATGGAGTGAGGACCCGAGATCAACAACATCACACCGCTGATCCAC 3378 			
Qy	21 SerleulleIgluJlUserGlnasnGlngInLylbsAsgngIngluleuLeuLeu 40 			
Dd	3379 AGCGTAGTCGAGGAGACCCAGAACCAACAGAAAGAACAGACAGAGACTGTCTGAGACTG 3438 			
Qy	41 AspLystrPalSerleutrTPAsnrTPPheAnlle 52 			
Dd	3439 GACACAGTGGCCACCGCTGTGGAACTGGTTCAACATC 3474 			
RESULT 10				
ABR91464				
ID	ABR91464 standard; DNA; 6505 BP.			
XX	ABR91464;			
AC				
XX				
DT	14-AUG-2002 (first entry)			
XX				
DE	Modified HIV protein-encoding plasmid DNA #16.			
XX				
KW	HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;			
KM	Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;			
KM	acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.			
XX				
OS	Synthetic.			
XX				
PN	WO200232943-A2.			
XX				
PD	25-APR-2002.			
PF	14-AUG-2001; 2001WO-US25721.			
XX				
PR	14-AUG-2000; 2000US-225097P.			
PR	14-NOV-2000; 2000US-252115P.			
PR	28-MAR-2001; 2001US-279257P.			
XX				
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.			
FA	(CHAD/) CHADRABARTI B K.			
XX				
PI	Nabel GJ, Huang Y;			
DR	WPI; 2002-452382/48.			
XX				
PT	New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and			
PT	its encoded protein, useful as vaccines for genetic or protein			
PT	immunisation for acquired immunodeficiency syndrome or human			
PT	immunodeficiency virus infection -			
PS	Disclosure; Page 411-413; 794pp: English.			
XX				
CC	The invention relates to a nucleic acid molecule encoding a modified HIV			
CC	(human immunodeficiency virus) protein. The nucleic acid molecule and its			
CC	encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for			
CC	genetic or protein immunisation to a host, respectively. In particular			
CC	these are useful for ameliorating the symptoms of acquired			
CC	immunodeficiency syndrome (AIDS) or HIV infection and generating an			
CC	antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,			
CC	Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent			
CC	compositions are useful for treating or preventing HIV infections or			
CC	AIDS. Sequences ABR91449-ABR91524 represent modified HIV protein-encoding			
CC	plasmid DNA sequences of the invention.			
XX				
SD	Sequence 6505 BP; 1637 A; 1769 C; 1644 G; 1455 T; 0 other;			



CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
CC compositions are useful for treating or preventing HIV infections or  
CC AIDS. Sequences ABR91449-ABR91624 represent modified HIV protein-encoding  
CC plasmid DNA sequences of the invention.

XX Sequence 6538 BP; 1624 A; 1771 C; 1674 G; 1469 T; 0 other;

Alignment Scores:

Pred. No.:	2,21e-28	Length:	6538
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	99.32%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-3 (1-52) x ABR91478 (1-6538)

QY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrHrSerLeuIleHis 20

DB 3339 AACCAACACCACTGGATGGAGTGGACCGCAGATCAACACTACACCACTGATCCAC 3398

QY 21 SerLeuIleGluGluSerGlnAsnGlnGluIleLysAsnGluGluLeuLeu 40

DB 3399 AGCTGATCGAGAGAGACCAACACAGAGAAGAAGACGAGAGCTGCTGAGCTG 3458

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52

DB 3459 GACAAAGTGGCCAGCTGTGGAACCTGTTCAACATC 3494

RESULT 13

ABR91481

ID ABR91481 standard; DNA: 6538 BP.

AC ABR91481;

XX 14-AUG-2002 (first entry)

DE Modified HIV protein-encoding plasmid DNA #33.

XX HIV; human immunodeficiency virus; gene: ds; circular; cyclic; anti-HIV;

KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KM acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

OS Synthetic.

XX WO200232943-A2.

PN 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

PR 14-NOV-2000; 2000US-252115P.

PR 28-MAR-2001; 2001US-279257P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA (CHAD/) CHADRABARTI B K.

XX Nabel GJ, Huang Y;

PI WPI; 2002-452382/48.

DR New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and

XX its encoded protein, useful as vaccines for genetic or protein

PT immunisation for acquired immunodeficiency syndrome or human

PT immunodeficiency virus infection -

XX Disclosure; Page 449-451; 794pp; English.

CC these are useful for ameliorating the symptoms of acquired  
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
CC compositions are useful for treating or preventing HIV infections or  
CC AIDS. Sequences ABR91449-ABR91624 represent modified HIV protein-encoding  
CC plasmid DNA sequences of the invention.

XX Sequence 6538 BP; 1627 A; 1764 C; 1677 G; 1470 T; 0 other;

Alignment Scores:

Pred. No.:	2,21e-28	Length:	6538
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	99.32%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-3 (1-52) x ABR91481 (1-6538)

QY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrHrSerLeuIleHis 20

DB 3339 AACCAACACCACTGGATGGAGTGGACCGCAGATCAACACTACACCACTGATCCAC 3398

QY 21 SerLeuIleGluGluSerGlnAsnGlnGluIleLysAsnGluGluLeuLeu 40

DB 3399 AGCTGATCGAGAGAGACCAACACAGAGAAGAAGACGAGAGCTGCTGAGCTG 3458

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52

DB 3459 GACAAAGTGGCCAGCTGTGGAACCTGTTCAACATC 3494

RESULT 14

ABR91479

ID ABR91479 standard; DNA: 6541 BP.

AC ABR91479;

XX 14-AUG-2002 (first entry)

DE Modified HIV protein-encoding plasmid DNA #31.

XX HIV; human immunodeficiency virus; gene: ds; circular; cyclic; anti-HIV;

KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KM acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

OS Synthetic.

XX WO200232943-A2.

PN 25-APR-2002.

XX 14-AUG-2001; 2001WO-US25721.

XX 14-AUG-2000; 2000US-225097P.

PR 14-NOV-2000; 2000US-252115P.

PR 28-MAR-2001; 2001US-279257P.

XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA (CHAD/) CHADRABARTI B K.

XX Nabel GJ, Huang Y;

PI WPI; 2002-452382/48.

DR New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and

XX its encoded protein, useful as vaccines for genetic or protein

PT immunisation for acquired immunodeficiency syndrome or human

PT immunodeficiency virus infection -

XX Disclosure; Page 445-447; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV



CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
 CC genetic or protein immunisation to a host, respectively. In particular  
 CC these are useful for ameliorating the symptoms of acquired  
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
 CC compositions are useful for treating or preventing HIV infections or  
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding  
 CC plasmid DNA sequences of the invention.

XX SQ Sequence 6541 BP; 1621 A; 1779 C; 1672 G; 1469 T; 0 other;

# Alignment Scores:

Pred. No.:	2,21e-28	Length:	6541
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	99.32%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-3 (1-52) x ABK91479 (1-6541)

QY 1 AsnHisThrThrPleuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
 DB 3342 AACACACACACCTGATGAGTGGACCGGAGATCAACAACTACACCGCTGATCCAC 3401

QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnIuLysAsnGluGlnIuLeuLeu 40  
 DB 3402 AGCCTATGAGAGAGAGCCAGAGACGAGAAAGAGAGAGAGCTGTGAGAGCTG 3461

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
 DB 3462 GACAAAGTGGGCCAGCTGTGGAAGTGTTCACATC 3497

RESULT 15

ABK91480  
 ID ABK91480 standard; DNA; 6577 BP.

XX AC ABK91480;

XX DT 14-AUG-2002 (first entry)

XX DE Modified HIV protein-encoding plasmid DNA #32.

XX KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;

XX KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

XX KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX OS Synthetic.

XX PN WO200232943-A2.

XX PD 25-APR-2002.

XX PF 14-AUG-2001; 2001WO-US25721.

XX PR 14-AUG-2000; 2000US-225097P.

XX PR 14-NOV-2000; 2000US-252115P.

XX PR 28-MAR-2001; 2001US-279257P.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PA (CHAD/) CHADRABARTI B K.

XX PI Nabel GJ, Huang Y;

XX DR WPI; 2002-452382/48.

XX PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and  
 PT its encoded protein, useful as vaccines for genetic or protein  
 PT immunisation for acquired immunodeficiency syndrome or human  
 PT immunodeficiency virus infection  
 XX

PS Disclosure; Page 447-449; 794pp; English.

XX CC The invention relates to a nucleic acid molecule encoding a modified HIV  
 CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
 CC genetic or protein immunisation to a host, respectively. In particular  
 CC these are useful for ameliorating the symptoms of acquired  
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
 CC compositions are useful for treating or preventing HIV infections or  
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding  
 CC plasmid DNA sequences of the invention.

XX SQ Sequence 6577 BP; 1635 A; 1782 C; 1680 G; 1480 T; 0 other;

# Alignment Scores:

Pred. No.:	2,23e-28	Length:	6577
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	99.32%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-3 (1-52) x ABK91480 (1-6577)

QY 1 AsnHisThrThrPleuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20

DB 3378 AACACACACACCTGATGAGTGGACCGGAGATCAACAACTACACCGCTGATCCAC 3437

QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnIuLysAsnGluGlnIuLeuLeu 40

DB 3438 AGCCTATGAGAGAGAGCCAGAGACGAGAAAGAGAGAGAGCTGTGAGAGCTG 3497

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52

DB 3498 GACAAAGTGGGCCAGCTGTGGAAGTGTTCACATC 3533

Search completed: March 25, 2003, 13:11:35  
 Job time : 207.877 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:42:33 ; Search time 39,9606 Seconds  
(without alignments)  
399.073 Million cell updates/sec

Title: US-09-877-606-3

Perfect score: 292

Sequence: 1 NHTTLEMDREINNTSLIH.....NEQELLEDKASLMMWNI 52

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh  
-O=/cgn2\_1/USPTO/spool/US09877606/runat\_18032003\_134801\_17715/app\_query.fasta\_1.1059  
-DB-issued\_Patents\_NA -QFMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi  
-LIST=45 -DOCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pic -NOR=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09877606.ecgn1.1.101@runat\_18032003\_134801\_17715 -NCPU=6 -ICPU=3  
-NO\_XLPPX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-NARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents NA:  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCRTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	99.3	1260	1	US-08-467-933-7
2	290	99.3	1568	1	US-08-589-446-7
3	290	99.3	1568	1	US-08-444-882-7
4	290	99.3	1568	1	US-08-389-459A-7
5	290	99.3	1568	1	US-08-987-867A-7
6	290	99.3	2571	2	US-07-916-098A-1
7	290	99.3	2694	1	US-08-147-890A-1
8	290	99.3	3084	1	US-08-147-890-2
9	290	99.3	699	5	PCT-US95-13335-2
10	277	94.9	1451	1	US-07-924-028A-1
11	277	94.9	1596	3	US-08-388-353-651
12	277	94.9	1596	3	US-08-488-551B-651

13	277	94.9	2644	4	US-08-472-240A-9	Sequence 9, Appl1
14	277	94.9	2730	3	US-08-728-122-1	Sequence 1, Appl1
15	277	94.9	2945	6	5462872-1	Patent No. 5462872
16	277	94.9	3563	3	US-08-463-210-6	Sequence 6, Appl1
17	277	94.9	7399	2	US-08-418-848A-9	Sequence 9, Appl1
18	277	94.9	8932	4	US-09-124-900-1	Sequence 1, Appl1
19	277	94.9	8933	3	US-08-463-210-4	Sequence 4, Appl1
20	277	94.9	8933	4	US-09-620-958A-3	Sequence 4, Appl1
21	277	94.9	8933	4	US-09-620-958A-4	Sequence 4, Appl1
22	277	94.9	8933	4	US-09-620-958A-9	Sequence 9, Appl1
23	277	94.9	9709	2	US-08-188-583-5	Sequence 5, Appl1
24	277	94.9	9709	3	US-08-388-353-1	Sequence 1, Appl1
25	277	94.9	9709	3	US-08-488-551B-1	Sequence 1, Appl1
26	277	94.9	9709	4	US-09-309-572-15	Sequence 15, Appl1
27	277	94.9	12494	4	US-08-935-312-13	Sequence 13, Appl1
28	277	94.9	12494	4	US-08-848-760B-33	Sequence 33, Appl1
29	277	94.9	15581	3	US-08-646-538-35	Sequence 35, Appl1
30	277	94.9	15581	4	US-09-503-222-35	Sequence 35, Appl1
31	272	93.8	2531	4	US-07-956-483-18	Sequence 18, Appl1
32	262	89.7	2184	4	US-09-337-387-4	Sequence 4, Appl1
33	259	88.7	3807	1	US-08-022-835-5	Sequence 5, Appl1
34	259	88.7	3807	1	US-08-388-809-5	Sequence 5, Appl1
35	259	88.7	3807	2	US-08-647-714-5	Sequence 5, Appl1
36	252	86.3	2552	2	US-08-448-603A-27	Sequence 27, Appl1
37	252	86.3	2552	3	US-09-134-075-27	Sequence 27, Appl1
38	252	86.3	2552	4	US-09-492-739-27	Sequence 27, Appl1
39	252	86.3	4527	2	US-08-944-449-8	Sequence 8, Appl1
40	252	86.3	4527	4	US-09-353-362-8	Sequence 8, Appl1
41	244	83.6	342	3	US-08-392-794A-5	Sequence 5, Appl1
42	244	83.6	345	3	US-08-392-794A-20	Sequence 20, Appl1
43	244	83.6	528	3	US-08-392-794A-7	Sequence 7, Appl1
44	244	83.6	531	3	US-08-392-794A-21	Sequence 21, Appl1
45	242	82.9	2571	1	US-08-254-358-3	Sequence 3, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-467-933-7  
Sequence 7, Application US/08467933  
Patent No. 5736391  
GENERAL INFORMATION:  
APPLICANT: Essex, Myron E.  
APPLICANT: Yu, Xiaofang  
APPLICANT: Lee, Tun-Hou  
TITLE OF INVENTION: GP41 MUTANTS AND THEIR USE AS  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
COMPUTER: IBM PS/2 Model 502 or 555X  
OPERATING SYSTEM: MS-DOS (Version 5.0)  
SOFTWARE: WordPerfect (Version 5.1)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/467,933  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/979,975  
FILING DATE: No. 5736391,ember 23, 1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Freeman, John W.  
REGISTRATION NUMBER: 29,066  
REFERENCE/DOCKET NUMBER: 00379/019001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070

```
;
; TELEFAX: (617) 542-8906
;
; TELEEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
;
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1260
;   TYPE: nucleic acid
;   STRANDEDNESS: double
;   TOPOLOGY: linear
;
US-08-467-933-7

Alignment Scores:
Pred. No.: 1.18e-29      Length: 1260
Score: 290.00           Matches: 51
Percent Similarity: 100.00%   Conservative: 1
Best Local Similarity: 98.08%   Mismatches: 0
Query Match: 99.32%          Indels: 0
DB: 1                    Gaps: 0

US-09-877-606-3 (1-52) x US-08-467-933-7 (1-1260)

QY 1 AsnHstHrThrTPlEuGlUTrPAsPaRgSlUleAsnAsnTyRThrSerLeuIleHis 20
    |||||||
Db 464 AATACACGACCTGGATGGAGTGGACAGAAATTAACATTACACAGCTTATATAC 523
    |||||||
QY 21 SerLeuIleGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuGluLeu 40
    |||||||
Db 524 TCCTTAATTGAAGATCGCAAAACGACGAAAGAAATGAACAATTAATTGGAATTA 583
    |||||||
QY 41 AsPlYsTrPAlaSerLeuTrPAsnTrPheAsnIle 52
    |||||||
Db 584 GATAAATGGCAAGTTGTGTGAATTGTTTAAACATA 619
    |||||||

RESULT 2
US-08-589-446-7
; Sequence 7, Application US/08589446
; Patent No. 5614413
;
; GENERAL INFORMATION:
;   APPLICANT: Morrow, Casey D.
;   TITLE OF INVENTION: ENCAPSITATED POLIOVIRUS NUCLEIC
;   TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: LAHIVE & COCKFIELD
;     STREET: 60 STATE STREET, SUITE 510
;     CITY: BOSTON
;     STATE: MASSACHUSETTS
;     COUNTRY: USA
;     ZIP: 02109
;
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/589,446
;   FILING DATE:
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US/08/087,009
;     FILING DATE:
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Geary III, William C.
;     REGISTRATION NUMBER: 31,359
;     REFERENCE/DOCKET NUMBER: UAG-004
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617) 227-7400
;     TELEFAX: (617) 227-5941
;
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1568 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     MOLECULE TYPE: single
```

```
;
; TOPOLOGY: linear
;
; MOLECULE TYPE: CDNA
;
; FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 7..1565
;
US-08-589-446-7

Alignment Scores:
Pred. No.: 1.55e-29      Length: 1568
Score: 290.00           Matches: 51
Percent Similarity: 100.00%   Conservative: 1
Best Local Similarity: 98.08%   Mismatches: 0
Query Match: 99.32%          Indels: 0
DB: 1                    Gaps: 0

US-09-877-606-3 (1-52) x US-08-589-446-7 (1-1568)

QY 1 AsnHstHrThrTPlEuGlUTrPAsPaRgSlUleAsnAsnTyRThrSerLeuIleHis 20
    |||||||
Db 1264 AATCACACGACCTGGATGGAGTGGACAGAAATTAACATTACACAGCTTATATAC 1323
    |||||||
QY 21 SerLeuIleGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuGluLeu 40
    |||||||
Db 1324 TCCTTAATTGAAGATCGCAAAACGACGAAAGAAATGAACAAGATTATGGAATTA 1383
    |||||||
QY 41 AsPlYsTrPAlaSerLeuTrPAsnTrPheAsnIle 52
    |||||||
Db 1384 GATAAATGGCAAGTTGTGTGAATTGTTTAAACATA 1419
    |||||||

RESULT 3
US-08-444-882-7
; Sequence 7, Application US/08444882
; Patent No. 5622705
;
; GENERAL INFORMATION:
;   APPLICANT: Morrow, Casey D.
;   TITLE OF INVENTION: ENCAPSITATED POLIOVIRUS NUCLEIC
;   TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
;   NUMBER OF SEQUENCES: 8
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: LAHIVE & COCKFIELD
;     STREET: 60 STATE STREET, SUITE 510
;     CITY: BOSTON
;     STATE: MASSACHUSETTS
;     COUNTRY: USA
;     ZIP: 02109
;
; COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/444,882
;   FILING DATE: 19-MAY-1995
;   CLASSIFICATION: 514
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER: US 08/087,009
;     FILING DATE: 01-JUL-1993
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Geary III, William C.
;     REGISTRATION NUMBER: 31,359
;     REFERENCE/DOCKET NUMBER: UAG-004
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (617) 227-7400
;     TELEFAX: (617) 227-5941
;
; INFORMATION FOR SEQ ID NO: 7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 1568 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: single
;     MOLECULE TYPE: CDNA
;     FEATURE:
```

```
; NAME/KEY: CDS
; LOCATION: 7..1565
US-08-444-882-7

Alignment Scores:
Pred. No.: 1.55e-29 Length: 1568
Score: 290.00 Matches: 51
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 0
Query Match: 99.32% Indels: 0
DB: 1 Gaps: 0

US-09-877-606-3 (1-52) x US-08-444-882-7 (1-1568)

QY 1 AsnHstHrThTrPlengLutRPaSPaRgGluIleAsnAsnTyRThSerleuIleHis 20
    |||||||
DB 1264 AATCAGACACCTGGATGGAGTGCGACAGAAATTAAACAATTACACAGCTTAATACAC 1323
QY 21 SerleuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeu 40
    |||||||
DB 1324 TCCTTAATTGAAGAAATCGCAAAACGACAGAAAGAAATGAACAGAAATTATTGGAATTA 1383
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
    |||||||
DB 1384 GATAAATGGCAAGTTTGTAATTGTTTAACATA 1419

RESULT 4
US-08-389-459A-7
; Sequence 7, Application US/08389459A
; Patent No. 5817512
; GENERAL INFORMATION:
; APPLICANT: MORROW, Casey D. and Porter, Donna, C.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT POLIOVIRUS
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/389,459A
; FILING DATE: 15-FEB-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,009
; FILING DATE: 01-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Silverl, Jean M.
; REGISTRATION NUMBER: 39,030
; REFERENCE/DOCKET NUMBER: UAG-004CP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1565
```

```
US-08-389-459A-7

Alignment Scores:
Pred. No.: 1.55e-29 Length: 1568
Score: 290.00 Matches: 51
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 0
Query Match: 99.32% Indels: 0
DB: 1 Gaps: 0

US-09-877-606-3 (1-52) x US-08-389-459A-7 (1-1568)

QY 1 AsnHstHrThTrPlengLutRPaSPaRgGluIleAsnAsnTyRThSerleuIleHis 20
    |||||||
DB 1264 AATCAGACACCTGGATGGAGTGCGACAGAAATTAAACAATTACACAGCTTAATACAC 1323
QY 21 SerleuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeu 40
    |||||||
DB 1324 TCCTTAATTGAAGAAATCGCAAAACGACAGAAAGAAATGAACAGAAATTATTGGAATTA 1383
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
    |||||||
DB 1384 GATAAATGGCAAGTTTGTAATTGTTTAACATA 1419

RESULT 5
US-08-987-867A-7
; Sequence 7, Application US/08987867A
; Patent No. 6063384
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/987,867A
; FILING DATE: 09-DEC-1997
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Myers, Louis
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: UAP-004CPDV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1565
US-08-987-867A-7

Alignment Scores:
```

Pred. No.: 1.55e-29 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 3 Gaps: 0

US-09-877-606-3 (1-52) x US-08-987-867A-7 (1-1568)

OY 1 AsnHsThrTrpLeuGluTrpAspArgGluLeuAsnAsnTyrrHsrLeuIleHis 20  
|||||  
Db 1264 AATACACGACCTGATGAGTAGTGAGACAGAAATTACATTACACAGCTTAATACAC 1323  
|||||  
OY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuLeu 40  
|||||  
Db 1324 TCCTTAATGAGATGCAAAACACAGAAAGATGAAATGAAATTATGGAATTA 1383  
|||||  
OY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATTAATGGCAAGTTGTGTGGAATTGGTTAACAATA 1419  
|||||

RESULT 6  
US-07-916-098A-1  
; Sequence 1, Application US/07916098A  
; Patent No. 5871732  
; GENERAL INFORMATION:  
; APPLICANT: BURKLY, LINDA C.  
; APPLICANT: CHISHOLM, PATRICIA L.  
; APPLICANT: THOMAS, DAVID W.  
; APPLICANT: ROSA, MARGARET D.  
; TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN  
; TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION  
; NUMBER OF SEQUENCES: 61  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.  
; STREET: 10 SOUTH WACKER DRIVE  
; CITY: CHICAGO  
; STATE: ILLINOIS  
; COUNTRY: U.S.A.  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WORD PERFECT 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/916,098A  
; FILING DATE: July 24, 1992  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/08843  
; FILING DATE: No. 5871732ember 27, 1991  
; CLASSIFICATION: 424  
; APPLICATION NUMBER: 07/618,542  
; FILING DATE: No. 5871732ember 27, 1990  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: JOHN J. MC DONNELL  
; REGISTRATION NUMBER: 26,949  
; REFERENCE/DOCKET NUMBER: 92,310-G  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312) 715-1000  
; TELEFAX: (312) 715-1234  
; TELEX: 910/221-5317  
; INFORMATION FOR SEO ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2571 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; HYPOTHETICAL: NO

ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 1..87  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 88..2568  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1533..1534  
OTHER INFORMATION: /note= "gp120/gp41 cleavage site"

US-07-916-098A-1  
OTHER INFORMATION: /note= "pre-HIV gp160"

Alignment Scores:  
Pred. No.: 2.88e-29 Length: 2571  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
DB: 2 Gaps: 0

US-09-877-606-3 (1-52) x US-07-916-098A-1 (1-2571)

OY 1 AsnHsThrTrpLeuGluTrpAspArgGluLeuAsnAsnTyrrHsrLeuIleHis 20  
|||||  
Db 1870 AATACACGACCTGATGAGTAGTGAGACAGAAATTACATTACACAGCTTAATACAC 1929  
|||||  
OY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuLeu 40  
|||||  
Db 1930 TCCTTAATGAGATGCAAAACACAGAAAGATGAAATGAAATTATGGAATTA 1989  
|||||

OY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1990 GATTAATGGCAAGTTGTGTGGAATTGGTTAACAATA 2025  
|||||

RESULT 7  
US-08-147-890-1  
; Sequence 1, Application US/08147890  
; Patent No. 5714316  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, David  
; APPLICANT: Williams, William  
; TITLE OF INVENTION: Chimeric Envelope Proteins for Viral  
; TITLE OF INVENTION: Targeting  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: Spring House Corporate Cntr, P.O. Box 457  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19477  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/147,890  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/7718,537  
; FILING DATE: 21-JUN-1991  
; ATTORNEY/AGENT INFORMATION:

NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST15AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2694 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
US-08-147-890-1

Alignment Scores:  
Pred. No.: 3.05e-29 Length: 2694  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
Gaps: 0  
DB: 1

US-09-877-606-3 (1-52) x US-08-147-890-1 (1-2694)

QY 1 AsnHstHrThrTPrLeuGlUTrpAspArgGluIleAsnAsnTyrrHrSerLeuIleHis 20  
DB 1978 AATCACACACCTCGATGGAGTGGGACAGACAGAAATTACACAGCTTAATACAC 2037

QY 21 SerLeuIleGluGlnSerGlnAsnGlnGlnGluLysAsnGluGlnGlnLeuLeuGluLeu 40  
DB 2038 TCCTTAATGGAAGATCGCAAAACCGCAAGAAAGATGACCAAGATATTGGAATTA 2097

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
DB 2098 GATAAATGGCAAGTTGTGGAATTGTTAACATA 2133

RESULT 8  
US-08-147-890-2  
Sequence 2, Application US/08147890  
Patent No. 5714316  
GENERAL INFORMATION:  
APPLICANT: Weiner, David  
APPLICANT: Williams, William  
APPLICANT: Levy, David N.  
TITLE OF INVENTION: Chimeric Envelope Proteins for Viral  
TITLE OF INVENTION: Targeting  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/147,890  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION NUMBER: US 07/718,537  
FILING DATE: 21-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST15AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3084 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
US-08-147-890-2

Alignment Scores:  
Pred. No.: 3.61e-29 Length: 3084  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 99.32% Indels: 0  
Gaps: 0  
DB: 1

US-09-877-606-3 (1-52) x US-08-147-890-2 (1-3084)

QY 1 AsnHstHrThrTPrLeuGlUTrpAspArgGluIleAsnAsnTyrrHrSerLeuIleHis 20  
DB 2368 AATCACACACCTCGATGGAGTGGGACAGAGAAATTACACAGCTTAATACAC 2427

QY 21 SerLeuIleGluGlnSerGlnAsnGlnGlnGluLysAsnGluGlnGlnLeuLeuGluLeu 40  
DB 2428 TCCTTAATGGAAGATCGCAAAACCGCAAGAAAGATGACCAAGATATTGGAATTA 2487

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
DB 2488 GATAAATGGCAAGTTGTGGAATTGTTAACATA 2523

RESULT 9  
PCT-US95-13335-2  
Sequence 2, Application PC/TUS9513335  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A RECOMBINANT PROTEIN USEFUL IN THE  
TITLE OF INVENTION: DETECTION OF HIV. DNA SEQUENCES ENCODING THE PROTEIN.  
TITLE OF INVENTION: IMMUNOSAAYS USING THE PROTEIN.  
NUMBER OF SEQUENCES: 2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13335  
FILING DATE:  
CLASSIFICATION:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 699 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
PCT-US95-13335-2

Alignment Scores:  
Pred. No.: 3.02e-28 Length: 699  
Score: 277.00 Matches: 49  
Percent Similarity: 98.08% Conservative: 2  
Best Local Similarity: 94.23% Mismatches: 1  
Query Match: 94.86% Indels: 0  
Gaps: 0  
DB: 5

US-09-877-606-3 (1-52) x PCT-US95-13335-2 (1-699)

QY 1 AsnHstHrThrTPrLeuGlUTrpAspArgGluIleAsnAsnTyrrHrSerLeuIleHis 20  
DB 388 AATAACATGACCTCGATGGAGTGGGACAGAGAAATTACCAATTACCAAGCTTAATACAT 447

QY 21 Serleuileglugluserglinsnglncglulysasnglucglulcgluleu 40  
|||||  
Db 448 TCCTTAATGGAAGATGCGCAACAGCAAGAAAGATGAACAATATTGGAATTA 507

QY 41 Asplystrpalaserleutripasntripheasnlle 52  
|||||  
Db 508 GATAAATGCGCAAGTTGTGGAATTTGTTTACATA 543

RESULT 10  
US-07-924-028A-1  
; Sequence 1, Application US/07924028A  
; Patent No. 5470573  
; GENERAL INFORMATION:  
; APPLICANT: Lubitz Werner, Szostak, Michael P.  
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS  
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Felfe & Lynch  
; STREET: 805 Third Avenue  
; CITY: New York City  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10022  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: PC-DOS  
; SOFTWARE: Wordperfect  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/924,028A  
; FILING DATE: 30-SEP-1992  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/EP91/00308  
; FILING DATE: 02-FEB-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 40 05 874  
; FILING DATE: 24-FEB-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Hanson, No. 5470573man D.  
; REGISTRATION NUMBER: 30,946  
; REFERENCE/DOCKET NUMBER: HUBR 1027  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 688-9200  
; TELEFAX: (212) 838-3884  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1451 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; US-07-924-028A-1

Alignment Scores:  
Pred. No.: 7.49e-28 Length: 1451  
Score: 277.00 Matches: 49  
Percent Similarity: 98.08% Conservative: 2  
Best Local Similarity: 94.23% Mismatches: 1  
Query Match: 94.86% Indels: 0  
DB: 1 Gaps: 0

US-09-877-606-3 (1-52) x US-07-924-028A-1 (1-1451)

QY 1 AsnHsthrtrtrPleuGlutRPaSPaRgLuIlleasnasnryrThSerleuIleHis 20  
|||||  
Db 496 AATAACATGACCTGATGCGATGGACAGACAAATTAACATTACACAGCTTAATAC 555

QY 21 Serleuileglugluserglinsnglncglulysasnglucglulcgluleu 40  
|||||  
Db 556 TCCTTAATGGAAGATGCGCAACAGCAAGAAAGATGAACAATATTGGAATTA 615

QY 41 Asplystrpalaserleutripasntripheasnlle 52

Db 616 GATAAATGCGCAAGTTGTGGAATTTGTTTACATA 651

RESULT 11  
US-08-388-353-651  
; Sequence 651, Application US/08388353  
; Patent No. 6010895  
; GENERAL INFORMATION:  
; APPLICANT: Deacon, Nicholas J.  
; APPLICANT: Learmont, Jennifer C.  
; APPLICANT: McPhee, Dale A.  
; APPLICANT: Crowe, Suzanne  
; APPLICANT: Cooper, David  
; TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
; NUMBER OF SEQUENCES: 800  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: United States  
; ZIP: 11530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/388,353  
; FILING DATE: 14-FEB-1995  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Digilio, Frank S.  
; REGISTRATION NUMBER: 31,346  
; REFERENCE/DOCKET NUMBER: 9606  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (516) 742-4343  
; TELEFAX: (516) 742-4366  
; TELEX: 230 901 SANS UR  
; INFORMATION FOR SEQ ID NO: 651:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1596 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; US-08-388-353-651

Alignment Scores:  
Pred. No.: 8.44e-28 Length: 1596  
Score: 277.00 Matches: 49  
Percent Similarity: 98.08% Conservative: 2  
Best Local Similarity: 94.23% Mismatches: 1  
Query Match: 94.86% Indels: 0  
DB: 3 Gaps: 0

US-09-877-606-3 (1-52) x US-08-388-353-651 (1-1596)

QY 1 AsnHsthrtrtrPleuGlutRPaSPaRgLuIlleasnasnryrThSerleuIleHis 20  
|||||  
Db 13 AATAACATGACCTGATGCGATGGACAGACAAATTAACATTACACAGCTTAATAC 72

QY 21 Serleuileglugluserglinsnglncglulysasnglucglulcgluleu 40  
|||||  
Db 73 TCCTTAATGGAAGATGCGCAACAGCAAGAAAGATGAACAATATTGGAATTA 132

QY 41 Asplystrpalaserleutripasntripheasnlle 52  
|||||  
Db 133 GATAAATGCGCAAGTTGTGGAATTTGTTTACATA 168

RESULT 12  
US-08-488-551B-651  
; Sequence 651, Application US/08488551B



[illegible]

```

: Patent No. 5284248
: GENERAL INFORMATION:
: APPLICANT: KIENY, Marie-Paule
: TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED
: TITLE OF INVENTION: 9p160 VARIANT
: NUMBER OF SEQUENCES: 29
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
: STREET: P.O. Box 1404
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/472,240A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/956,483
: FILING DATE: 31-DEC-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Teskin, Robin L.
: REGISTRATION NUMBER: 35,030
: REFERENCE/DOCKET NUMBER: 017753-055
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2644 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: not relevant
: TOPOLOGY: not relevant
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 12..2600
: US-08-472-240A-9
:
: Alignment Scores:
: Pred. No.: 1.58e-27 Length: 2644
: Score: 277.00 Matches: 49
: Percent Similarity: 98.08% Conservative: 2
: Best Local Similarity: 94.23% Mismatches: 1
: Query Match: 94.86% Indels: 0
: DB: 4 Gaps: 0
:
: US-09-877-606-3 (1-52) x US-08-472-240A-9 (1-2644)
:
: QY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20
: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
: Db 1896 AATAACATGACCTGAGATGAGATGGAGACAGAGAAATTAAACAATTACACAAGCTTAAATACAT 1955
:
: QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluGluGluGluGluGluGluGluGluGlu 40
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
: Db 1956 TCCCTTAATTGAAAGATTCGCAAAACGACGCAAGAAAGAAAGAAAGAAAGAAATTAATGGAATTA 2015
:
: QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
: ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
: Db 2016 GATTAATGGCGCAACTTTGTGCAATTGCTTAACATA 2051
:
: RESULT 14
: US-08-728-122-1
: ; Sequence 1, Application US/08728122
: ; Patent No. 6074646
: ; GENERAL INFORMATION:
: ; APPLICANT: Cloyd, Miles W.
: ; APPLICANT: Ramsey, Keith M.

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 15:49:42 ; Search time 73.5172 Seconds  
(without alignments)  
548.507 Million cell updates/sec

Title: US-09-877-606-3  
Perfect score: 292  
Sequence: 1 NHTTWLEMDREINNTSLIH.....NEOELLEDKWASLWMMENI 52

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 538826 segs, 387737923 residues  
Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+12n.model -DEV=xlh  
-O=/cgn2.1/USPTO/spool/US0987766/runat\_18032003.134802.17783/app\_query.fasta.1.1059  
-DB=Published\_Applications\_NA -QFMT=fastlap -SUFFIX=rmpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pcpt -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US0987766.ecgn.1.1.115.6runat.18032003.134802.17783  
-NCPU=6 -ICPU=3 -NO\_XLPHY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications\_NA:\*

- 1: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2.6/ptodata/2/pubpna/PC1\_NEW\_PUB.seq:\*
- 3: /cgn2.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2.6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2.6/ptodata/2/pubpna/PC1US\_PUBCOMB.seq:\*
- 7: /cgn2.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2.6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 10: /cgn2.6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*
- 12: /cgn2.6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*
- 13: /cgn2.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 14: /cgn2.6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	290	99.3	1568 10	US-09-756-551A-7
2	290	99.3	9719 10	US-09-737-190A-1
3	277	94.9	2280 9	US-10-003-035-16
4	277	94.9	2562 9	US-10-003-035-32

5	277	94.9	2583 9	US-10-003-035-24	Sequence 24, Appl
6	277	94.9	2746 9	US-10-003-035-20	Sequence 20, Appl
7	277	94.9	2747 9	US-10-003-035-23	Sequence 23, Appl
8	277	94.9	2941 9	US-10-003-035-19	Sequence 19, Appl
9	277	94.9	2950 9	US-10-003-035-22	Sequence 22, Appl
10	277	94.9	3157 9	US-10-003-035-14	Sequence 14, Appl
11	277	94.9	3417 9	US-10-003-035-21	Sequence 21, Appl
12	277	94.9	3839 9	US-10-003-035-52	Sequence 52, Appl
13	277	94.9	4040 9	US-10-003-035-54	Sequence 54, Appl
14	277	94.9	8933 10	US-09-943-286-3	Sequence 3, Appl
15	277	94.9	8933 10	US-09-943-286-4	Sequence 4, Appl
16	277	94.9	8933 10	US-09-943-286-9	Sequence 9, Appl
17	277	94.9	9918 10	US-09-798-675-5	Sequence 5, Appl
18	251	86.0	1860 10	US-09-891-609-3	Sequence 1, Appl
19	251	86.0	1941 10	US-09-891-609-1	Sequence 1, Appl
20	251	86.0	2298 10	US-09-476-242-22	Sequence 22, Appl
21	251	86.0	2298 10	US-09-476-242-23	Sequence 23, Appl
22	251	86.0	2298 10	US-09-476-242-24	Sequence 24, Appl
23	251	86.0	2310 10	US-09-476-242-3	Sequence 3, Appl
24	251	86.0	2310 10	US-09-476-242-21	Sequence 21, Appl
25	251	86.0	2316 10	US-09-476-242-4	Sequence 4, Appl
26	251	86.0	2316 10	US-09-476-242-8	Sequence 8, Appl
27	251	86.0	2322 10	US-09-476-242-5	Sequence 5, Appl
28	251	86.0	2322 10	US-09-476-242-18	Sequence 18, Appl
29	251	86.0	2322 10	US-09-476-242-19	Sequence 19, Appl
30	251	86.0	2322 10	US-09-476-242-20	Sequence 20, Appl
31	251	86.0	2328 10	US-09-476-242-6	Sequence 6, Appl
32	251	86.0	2334 10	US-09-476-242-7	Sequence 7, Appl
33	251	86.0	2352 10	US-09-476-242-26	Sequence 26, Appl
34	251	86.0	2358 10	US-09-476-242-25	Sequence 25, Appl
35	251	86.0	2517 10	US-09-476-242-16	Sequence 16, Appl
36	251	86.0	2517 10	US-09-476-242-17	Sequence 17, Appl
37	251	86.0	2523 10	US-09-476-242-15	Sequence 15, Appl
38	251	86.0	2529 10	US-09-476-242-14	Sequence 14, Appl
39	251	86.0	2535 10	US-09-476-242-13	Sequence 13, Appl
40	251	86.0	2541 10	US-09-476-242-9	Sequence 9, Appl
41	251	86.0	2541 10	US-09-476-242-10	Sequence 10, Appl
42	251	86.0	2541 10	US-09-476-242-11	Sequence 11, Appl
43	251	86.0	2541 10	US-09-476-242-12	Sequence 12, Appl
44	242	82.9	2571 9	US-10-077-294-3	Sequence 3, Appl
45	242	82.9	2571 9	US-10-163-886-3	Sequence 3, Appl

## ALIGNMENTS

RESULT 1  
US-09-756-551A-7  
; Sequence 7, Application US/09756551A  
; Patent No. US20020051768A1  
; GENERAL INFORMATION:  
; APPLICANT: C. MORROW et al  
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL  
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 STATE STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/756,551A  
; FILING DATE: 08-JAN-2001  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/376,184

Alignment Scores:		
Pred. No.:	7,09e-30	Length: 9719
Score:	230.00	Matches: 51
Percent Similarity:	100.00%	Conservative: 1
Best Local Similarity:	98.08%	Mismatches: 0
Query Match:	99.32%	Indels: 0
DB:	10	Gaps: 0
US-09-877-606-3 (1-52) x US-09-737-190A-1 (1-9719)		
QY	1	ASnHsHrHrTrPLeuGlUTrPAsPArYgGluIlLeAsAsnTyrrThrSerLeuIlLeHs 20
DB	8094	AATACACGACCGATGGATGGAGTGGACAGAAATTAACAAATTACACAAGCTTAATACAC 8153
QY	21	SerLeuIlLeGlUuSerGlnAsnGlnGlnGluLyAsnGluGlnGluLeu 40
DB	8154	TCCTATTATGACAAATCGCAAAACGACGAAGAAAGAAATGAACAAAGATTATTGGAATTA 8213
QY	41	AspLystrPalasSerLeuTrpAsnTrpPheAsnIlle 52
DB	8214	GATMAATGGCAAGTTTGATGGAAATGTGTTTAACATA 8249
RESULT 3		
US-10-003-035-16		
Sequence 16, Application US/10003035		
Patent No. US20020155127A1		
GENERAL INFORMATION:		
APPLICANT: Wang, Danher		
TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS		
FILE REFERENCE: 22488-712		
CURRENT APPLICATION NUMBER: US/10/003,035		
PRIOR APPLICATION NUMBER: 2001-11-01		
PRIOR FILING DATE: 2000-06-02		
PRIOR APPLICATION NUMBER: PCT/US01/18238		
PRIOR FILING DATE: 2001-06-04		
NUMBER OF SEQ. ID NOS: 75		
SOFTWARE: PatentIn version 3.1		
SEQ ID NO 16		
LENGTH: 2280		
TYPE: DNA		
ORGANISM: Artificial sequence		
FEATURE:		
OTHER INFORMATION: Modified Env		
US-10-003-035-16		
Alignment Scores:		
Pred. No.:	6.62e-29	Length: 2280
Score:	277.00	Matches: 49
Percent Similarity:	98.08%	Conservative: 2
Best Local Similarity:	94.23%	Mismatches: 1
Query Match:	94.86%	Indels: 0
DB:	9	Gaps: 0
US-09-877-606-3 (1-52) x US-10-003-035-16 (1-2280)		
QY	1	ASnHsHrHrTrPLeuGlUTrPAsPArYgGluIlLeAsAsnTyrrThrSerLeuIlLeHs 20
DB	1876	AATACACGACCGATGGATGGAGTGGACAGAAATTAACAAATTACACAAGCTTAATACAC 19353
QY	21	SerLeuIlLeGlUuSerGlnAsnGlnGlnGluLyAsnGluGlnGluLeu 40
DB	1936	TCCTATTATGACAAATCGCAAAACGACGAAGAAAGAAATGAACAAAGATTATTGGAATTA 19955
QY	41	AspLystrPalasSerLeuTrpAsnTrpPheAsnIlle 52
DB	1996	GATMAATGGCAAGTTTGATGGAAATGTGTTTAACATA 2031
RESULT 4		
US-10-003-035-32		
Sequence 32, Application US/10003035		
Patent No. US20020155127A1		
GENERAL INFORMATION:		

PRIOR APPLICATION NUMBER: PCT/US01/18238

```

      : PRIOR FILING DATE: 2001-06-04
      : NUMBER OF SEQ ID NOS: 75
      : SOFTWARE: PatentIn version 3.1
      : SEQ ID NO 23
      : LENGTH: 2747
      : TYPE: DNA
      : ORGANISM: Artificial sequence
      : FEATURE:
      : OTHER INFORMATION: Modified Env/Tat from strain BH10
US-10-003-035-23

Alignment Scores:
Pred. No.:
Score: 8.47e-29          Length: 2747
Percent Similarity: 277.00    Matches: 49
Best Local Similarity: 98.088   Conservative: 2
Query Match: 94.238           Mismatches: 1
DB: 94.86%                   Indels: 0
                                Gaps: 0

US-09-877-606-3 (1-52) x US-10-003-035-23 (1-2747)

OY      1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTYrThSerleuIleHis 20
Db      2343 AATPACATGCCTGGATGGAGTGTGGACAGAAATTAAACAATTACACAGCTTAATACAC 2402
OY      21 SerleuIleGluIserGlnAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40
Db      2403 TCCTTAATTGAAGAATCGCAAAACCAGCAAGAAAGAAATGAACAGAAATTTATGGAAATTA 2462
OY      41 AspIlyStrPalaseSerleuTrpAsnTrpPheasnille 52
Db      2463 GATAAATGGCAGATTGTGTGCAATTGCTTTAACATA 2498

RESULT 8
US-10-003-035-19
: Sequence 19, Application US/10003035
: Patent No. US20020155127A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Danher
: TITLE OR INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
: FILE REFERENCE: 22488-712
: CURRENT APPLICATION NUMBER: US/10/003,035
: PRIOR FILING DATE: 2001-11-01
: PRIOR APPLICATION NUMBER: 09/585,599
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US01/18238
: PRIOR FILING DATE: 2001-06-04
: NUMBER OF SEQ ID NOS: 75
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 19
: LENGTH: 2941
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: Modified Env from HIV strain pNL4-3
US-10-003-035-19

Alignment Scores:
Pred. No.:
Score: 9.26e-29          Length: 2941
Percent Similarity: 277.00    Matches: 49
Best Local Similarity: 98.088   Conservative: 2
Query Match: 94.238           Mismatches: 1
DB: 94.86%                   Indels: 0
                                Gaps: 0

US-09-877-606-3 (1-52) x US-10-003-035-19 (1-2941)

OY      1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTYrThSerleuIleHis 20
Db      2336 AATPACATGCCTGGATGGAGTGTGGACAGAAATTAAACAATTACACAGCTTAATACAC 2395
OY      21 SerleuIleGluIserGlnAsnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 40

```

```

Db 2296 TCCATTATGAGAAATCGCAAAACCGACGAAGAAAGAAATGACCAAGATTATTGAATTA 2455
OY 41 Asplystirpalaserleuttripsanttripleasnlle 52
|||||
Db 2456 GATTAATGGCGCAAGTTTGTCGGAATTTGTTTAACATA 2491
|||||

RESULT 9
US-10-003-035-22
; Sequence 22, Application US/10003035
; Patent No. US2002015127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 2950
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/NeF from strain BH10
US-10-003-035-22

Alignment Scores:
Pred. No.: 9, 3e-29 Length: 2950
Score: 277.00 Matches: 49
Percent Similarity: 98.08% Conservative: 2
Best Local Similarity: 94.23% Mismatches: 1
Query Match: 94.86% Indels: 0
Gaps: 0

US-09-877-606-3 (1-52) x US-10-003-035-22 (1-2950)
OY 1 AsnHisThrThrTrypLeuGluTrpAspArgGluIleasnAsnTyrThrSerLeuIleHis 20
|||||
Db 1876 AATRAACTGACCTGGATGGAGTGGGCGACAGAAATTAACAACTTAAATACAC 1935
|||||
OY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuLeuGluLeu 40
|||||
Db 1936 TCCATTATGCAAGATTCGCAAAACCGACGAAGAAAGAAATGACCAAGATTATTGAATTA 1995
|||||
OY 41 Asplystirpalaserleuttripsanttripleasnlle 52
|||||
Db 1996 GATTAATGGCGCAAGTTTGTCGGAATTTGTTTAACATA 2031
|||||

RESULT 10
US-10-003-035-14
; Sequence 14, Application US/10003035
; Patent No. US2002015127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 3157
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:

```

```

Sequence 52, Application US/10003035
Patent No. US20020155127A1
GENERAL INFORMATION:
APPLICANT: Wang, Danher
TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
FILE REFERENCE: 22488-712
CURRENT APPLICATION NUMBER: US/10/003,035
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: 09/585,599
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US01/18238
PRIOR FILING DATE: 2001-06-04
NUMBER OF SEQ. ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 52
LENGTH: 3839
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Modified Env/Tat/Rev
US-10-003-035-52

Alignment Scores:
Pred. No.: 1,32e-28 Length: 3839
Score: 277.00 Matches: 49
Percent Similarity: 98.088 Conservative: 2
Best Local Similarity: 94.238 Mismatches: 1
Query Match: 94.866 Indels: 0
DB: Gaps: 0

US-09-877-606-3 (1-52) x US-10-003-035-52 (1-3839)
QY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20
Db 3435 AATAACATGACCTGCGATGAGTGTGGACAGAGAATTAACAATTACACAAGCTTAATACAC 3494
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluIleAsnGluGluLeuLeu 40
Db 3495 TCCTTAATTGAAGATGCGAACAACCGACAAGAAAGAAATGAACAGAAATTATTGGAATTA 3554
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
Db 3555 GATAAATGGCGAAGTTGTGTGAATTGGTTTAAACATA 3590

RESULT 13
US-10-003-035-54
Sequence 54, Application US/10003035
Patent No. US20020155127A1
GENERAL INFORMATION:
APPLICANT: Wang, Danher
TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
FILE REFERENCE: 22488-712
CURRENT APPLICATION NUMBER: US/10/003,035
PRIOR FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: 09/585,599
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US01/18238
PRIOR FILING DATE: 2001-06-04
NUMBER OF SEQ. ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 54
LENGTH: 4040
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Modified Env/Tat/Rev
US-10-003-035-54

Alignment Scores:
Pred. No.: 1.41e-28 Length: 4040
Score: 277.00 Matches: 49
Percent Similarity: 98.088 Conservative: 2
Best Local Similarity: 94.238 Mismatches: 1
DB: Mismatches: 1

```





GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: March 18, 2003, 14:30:36 ; Search time 16.6502 Seconds  
(without alignments)  
300.235 Million cell updates/sec

Title: US-09-877-606-3

Perfect score: 292

Sequence: 1 NHTTWLEMDREINNTSLIH.....NEQELLELDKMASLNMWNI 52

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: pir1:\*  
3: pir2:\*  
4: pir3:\*  
5: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	277	94.9	851 2	S33985 env polyprotein -
2	277	94.9	854 2	S13288 env protein - huma
3	277	94.9	856 1	VCLJH3 env polyprotein pr
4	277	94.9	861 1	VCLJLV env polyprotein pr
5	268	91.8	856 1	VCLJLV env polyprotein pr
6	260	89.0	358 2	S21998 envelope protein g
7	254	87.0	856 1	VCLJ3W env polyprotein pr
8	253	86.6	861 1	VCLJSC env polyprotein pr
9	250	85.6	443 2	C41621 env polyprotein p
10	247	84.6	847 2	T09448 envelope glycoprot
11	247	84.6	847 2	S13289 env protein - huma
12	245	83.9	852 2	T12016 envelope glycoprot
13	244	83.6	357 2	S21996 envelope protein g
14	244	83.6	358 2	S22002 envelope protein g
15	244	83.6	358 2	S22000 envelope protein g
16	242	82.9	358 2	S70417 envelope protein g
17	242	82.9	859 1	VCLJMN env polyprotein pr
18	241	82.5	357 2	S22006 envelope protein g
19	241	82.5	357 2	S21994 envelope protein g
20	241	82.5	853 2	S54384 envelope polyprote
21	241	82.5	853 1	VCLJZR env polyprotein pr
22	240	82.2	357 2	S21992 envelope protein g
23	238	81.5	852 2	VCLJBR env polyprotein -
24	237	81.2	445 2	A41621 env polyprotein M
25	236	80.8	357 2	S22004 envelope protein g
26	236	80.8	729 1	VCLJXK env polyprotein pr
27	236	80.8	855 1	VCLJAJ env polyprotein pr
28	236	80.8	861 1	VCLJKB env polyprotein pr
29	234	80.1	846 1	VCLJND env polyprotein pr

30	233	79.8	868 1	VCLJH4 env polyprotein -
31	232	79.5	843 1	H44001 env polyprotein pr
32	229	78.4	859 2	T01672 envelope polyprote
33	228	78.1	136 2	JU0266 envelope polyprote
34	227	77.7	454 2	B41621 env polyprotein D
35	223	76.4	136 2	JT0954 envelope polyprote
36	213	72.9	357 2	S21990 envelope protein g
37	213	72.9	856 1	A44963 env polyprotein pr
38	198	67.8	854 1	VCLJST env polyprotein pr
39	192	65.8	877 2	S49197 envelope protein p
40	177	60.6	863 2	A53034 gag polyprotein -
41	130	44.5	151 2	S30448 env protein - huma
42	130	44.5	151 2	S30452 env protein - huma
43	130	44.5	151 2	S30451 env protein - huma
44	130	44.5	881 1	VCLJG3 env polyprotein -
45	130	44.5	885 2	S04322 env polyprotein -

#### ALIGNMENTS

##### RESULT 1

S33985 env polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999

C:Accession: S33985

R:Carlini, F.

submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33985

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-851 <CAR>

A:Cross-references: EMBL:Z11530; NID:960192; PIDN:CAA77628.1; PID:960199

C:Superfamily: type E retrovirus env polyprotein

Query Match

Best Local Similarity 94.9%; Score 277; DB 2; Length 851;

Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NHTTWLEMDREINNTSLIHSLIEESQOQEKNEQELLELDKMASLNMWNI 52

Db 619 NNMTWMEMDREINNTSLIHSLIEESQOQEKNEQELLELDKMASLNMWNI 670

##### RESULT 2

S13288 env protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997

C:Accession: S13288

R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, Nature 348, 69-73, 1990

A:title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp1

A:Reference number: S13288; MUID:91043044; PMID:2172833

A:Accession: S13288

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-854 <OBR>

C:Superfamily: type E retrovirus env polyprotein

Query Match

Best Local Similarity 94.9%; Score 277; DB 2; Length 854;

Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NHTTWLEMDREINNTSLIHSLIEESQOQEKNEQELLELDKMASLNMWNI 52

Db 622 NNMTWMEMDREINNTSLIHSLIEESQOQEKNEQELLELDKMASLNMWNI 673

##### RESULT 3

VCLJH3

env polyprotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_rev10n 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03973  
R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Doranberger, J.A.; Papas, T.S.; Ghayab, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.  
Nature 313, 277-284, 1985  
A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.  
A:Reference number: A93353; MUID:85111123; PMID:2578615  
A:Accession: A03973  
A:Molecule type: DNA  
A:Residues: 1-856 <RAT>  
A:Cross-references: GB:M1564; GB:K02008; GB:K02009; GB:K02010; NID:g326383; PIDN:AAA442  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS: capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,241,262,276,289,295,301,332,339,356,386,392,397,406  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic  
Query Match 94.9%; Score 277; DB 1; Length 856;  
Best Local Similarity 94.2%; Pred. No. 6.5e-22;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 1 NHTTWLEMDREINNTSLIHSLEESONQOEKNEDELLELDKWSLWNNFNI 52  
Db 624 NNMWTWEMDREINNTSLIHSLEESONQOEKNEDELLELDKWSLWNNFNI 675

RESULT 4  
VCLJLV  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_rev10n 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03975  
R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.  
Cell 40, 9-17, 1985  
A:Title: Nucleotide sequence of the AIDS virus, LAV.  
A:Reference number: A90866; MUID:85099333; PMID:2981635  
A:Accession: A03975  
A:Molecule type: DNA  
A:Residues: 1-861 <NAL>  
A:Cross-references: GB:K02013; NID:g326417; PIDN:AAB59751.1; PID:g326424  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS: capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:517-861/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,146,161,165,191,204,239,246,267,281,294,300,306,337,344,361,391,397,402,411  
F:616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predic  
Query Match 94.9%; Score 277; DB 1; Length 861;  
Best Local Similarity 94.2%; Pred. No. 6.5e-22;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 1 NHTTWLEMDREINNTSLIHSLEESONQOEKNEDELLELDKWSLWNNFNI 52  
Db 629 NNMWTWEMDREINNTSLIHSLEESONQOEKNEDELLELDKWSLWNNFNI 680

RESULT 5  
VCLJLV  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)  
N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_rev10n 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03974  
R:Muesing, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.  
Nature 313, 450-458, 1985  
A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retr  
A:Reference number: A93355; MUID:85111157; PMID:2982104  
A:Accession: A03974  
A:Molecule type: DNA  
A:Residues: 1-856 <MDP>  
A:Cross-references: GB:K02083; NID:g555008; PIDN:AAB59873.1; PID:g328559  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS: capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,241,262,276,289,295,301,332,339,356,386,392,397,  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status pre  
Query Match 91.8%; Score 268; DB 1; Length 856;  
Best Local Similarity 90.4%; Pred. No. 6e-21;  
Matches 47; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
OY 1 NHTTWLEMDREINNTSLIHSLEESONQOEKNEDELLELDKWSLWNNFNI 52  
Db 624 NNMWTWEMDREINNTSLIHSLEESONQOEKNEDELLELDKWSLWNNFNI 675

RESULT 6  
S21998  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 28  
C:Date: 20-Feb-1995 #sequence\_rev10n 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S21998; S70425  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
Submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as dete  
A:Reference number: S21990  
A:Accession: S21998  
A:Molecule type: DNA  
A:Residues: 1-358 <STEL>  
A:Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43630.1; PID:g60183  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
Aids Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer  
A:Reference number: S70417; MUID:92144205; PMID:1736940  
A:Accession: S70425  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-222, 'X', 224-358 <STE2>  
A:Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43630.1; PID:g60183  
C:Superfamily: type E retrovirus env polyprotein  
Query Match 89.0%; Score 260; DB 2; Length 358;  
Best Local Similarity 86.5%; Pred. No. 1.6e-20;  
Matches 45; Conservative 6; Mismatches 1; Indels 0; Gaps 0;  
OY 1 NHTTWLEMDREINNTSLIHSLEESONQOEKNEDELLELDKWSLWNNFNI 52  
Db 126 NNMWTWEMDREINNTSLIYLTLEESONQOEKNEDELLELDKWSLWNNFNI 177

RESULT 7  
VCLJ3W  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate WKJ1)  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 30-Sep-1988 #sequence\_rev10n 30-Sep-1988 #text\_change 07-Nov-1997  
C:Accession: A24774

R:Starcich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.  
Cell 45, 637-648, 1986  
A:Title: Identification and characterization of conserved and variable regions in the en  
A:Reference number: A24774; M0ID:86218077; PMID:2423250  
A:Accession: A24774  
A:Molecule type: DNA  
A:Residues: 1-856 <STA>  
A:Cross-references: GB:K03455; GB:M38432; NID:91906382  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-501/Product: coat protein gp120 #status predicted <GP1>  
F:502-847/Product: coat protein gp41 #status predicted <GP2>  
F:87,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,455

Query Match 87.0%; Score 254; DB 1; Length 856;  
Best Local Similarity 84.6%; Pred. No. 1,9e-19;  
Matches 44; Conservative 7; Mismatches 1; Indels 0; Gaps 0;

Oy 1 NHTTLEMDREINNTSLIHSLEESONQOEKNEQELLELDKWSLNMWFI 52  
Db 624 NNMWTMWEREIDNYTSLIYLIIESONQOEKNEQELLELDKWSLNMWFSI 675

RESULT 8  
VCLISC  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997  
C:Accession: B28922  
R:Curcio, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta  
Virology 164, 531-536, 1988  
A:Title: Envelope sequences of two new United States HIV-1 isolates.  
A:Reference number: A28922; M0ID:88219542; PMID:3369091  
A:Accession: B28922  
A:Molecule type: DNA  
A:Residues: 1-861 <CUR>  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote  
F:1-29/Domain: signal sequence #status predicted <SIG>  
F:30-861/Product: env polyprotein #status predicted <GP>  
F:87,129,135,140,143,159,187,198,234,241,262,276,295,301,302,333,340,356,362,386,396

Query Match 86.6%; Score 253; DB 1; Length 861;  
Best Local Similarity 89.8%; Pred. No. 2,5e-19;  
Matches 44; Conservative 5; Mismatches 0; Indels 0; Gaps 0;

Oy 4 TWLEMDREINNTSLIHSLEESONQOEKNEQELLELDKWSLNMWFI 52  
Db 632 TMMWEMEREIDNYTSLIYLIIESONQOEKNEQELLELDKWSLNMWFI 680

RESULT 9  
C41621  
env polyprotein p - human immunodeficiency virus type 1 (fragment)  
N:Alternate names: coat polyprotein  
C:Species: amino end of coat protein gp41; carboxyl end of coat protein gp120  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 11-Feb-1993 #sequence\_revision 31-Dec-1993 #text\_change 26-Aug-1999  
C:Accession: C41621  
R:Burgger, H.; Weiser, B.; Flaeherty, K.; Guila, J.; Nguyen, P.N.; Gibbs, R.A.  
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991  
A:Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity  
A:Reference number: A41621; M0ID:92107924; PMID:1763038  
A:Accession: C41621  
A:Molecule type: DNA

A:Residues: 1-443 <BUR>  
A:Cross-references: GB:M77230; NID:9328631; PIDN:AB03792.1; PID:9555015  
A:Note: this virus was isolated from the mother's sexual partner  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp  
F:1-251/Product: coat protein gp120 (fragment) #status predicted <GP1>  
F:252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>  
F:424-443/Domain: transmembrane #status predicted <TM>  
F:9,23,36,48,78,101,107,131,143,147,153,188,200,203,351,356,365,377/Binding site:

Query Match 85.6%; Score 250; DB 2; Length 443;  
Best Local Similarity 82.7%; Pred. No. 2,4e-19;  
Matches 43; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NHTTLEMDREINNTSLIHSLEESONQOEKNEQELLELDKWSLNMWFI 52  
Db 364 NNMWTMWEREIDNYTSLIYLIIESONQOEKNEQELLELDKWSLNMWFSI 415

RESULT 10  
T09448  
envelope glycoprotein - human immunodeficiency virus type 1 (strain JREL)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999  
C:Accession: T09448  
R:Pang, S.; Vinters, H.V.; Akshl, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namaz  
submitted to the EMBL Data Library, July 1996  
A:Reference number: Z16673  
A:Accession: T09448  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-847 <PAN>  
A:Cross-references: EMBL:U63632; NID:91465777; PID:91465781  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein

Query Match 84.6%; Score 247; DB 2; Length 847;  
Best Local Similarity 82.7%; Pred. No. 1,1e-18;  
Matches 43; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NHTTLEMDREINNTSLIHSLEESONQOEKNEQELLELDKWSLNMWFI 52  
Db 615 NNMWTMWEREIDNYTSLIYLIIESONQOEKNEQELLELDKWSLNMWFI 666

RESULT 11  
S13289  
env protein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C:Accession: S13289  
R:O'Brien, W.A.; Koyanagi, Y.; Namaz, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack,  
Nature 348, 69-73, 1990  
A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp1  
A:Reference number: S13288; M0ID:91043044; PMID:2172833  
A:Accession: S13289  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-847 <OBR>  
C:Superfamily: type E retrovirus env polyprotein

Query Match 84.6%; Score 247; DB 2; Length 847;  
Best Local Similarity 82.7%; Pred. No. 1,1e-18;  
Matches 43; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

Oy 1 NHTTLEMDREINNTSLIHSLEESONQOEKNEQELLELDKWSLNMWFI 52  
Db 615 NNMWTMWEREIDNYTSLIYLIIESONQOEKNEQELLELDKWSLNMWFI 666

RESULT 12  
T12016  
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T12016  
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.  
AIDS Res. Hum. Retroviruses 14, 329-337, 1998  
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in S  
A:Reference number: 217379; MUID:98178716; PMID:9518984  
A:Accession: T12016  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-852 <MCC>  
A:Cross-references: EMBL:U90934; NID:92351783; PIDN:AAC59271.1; PID:92351784  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein

Query Match 83.6%; Score 244; DB 2; Length 852;  
Best Local Similarity 82.7%; Pred. No. 1.8e-18;  
Matches 43; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWFWNI 52  
Db 620 NMTWMEWDERIHNYTSLIYTLIESONQOEKNEDELLELDKMSLWFWNI 671

RESULT 13  
S21996  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)

C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S70422; S21996  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr  
A:Reference number: S70417; MUID:92144209; PMID:11736940  
A:Accession: S70422  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-357 <ST2>  
A:Cross-references: EMBL:X61356; NID:960181; PIDN:CAA3624.1; PID:q1067129  
A:Experimental source: patient 27L  
A:Note: submitted to the EMBL Data Library, July 1991  
C:Superfamily: type E retrovirus env polyprotein

Query Match 83.6%; Score 244; DB 2; Length 357;  
Best Local Similarity 80.8%; Pred. No. 8.1e-19;  
Matches 42; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWFWNI 52  
Db 125 DNMTWMEWDERIHNYTSLIYTLIESONQOEKNEDELLELDKMSLWFWNI 176

RESULT 14  
S22002

envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 3L  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S22002; S70418  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determ

A:Reference number: S21990  
A:Accession: S22002  
A:Molecule type: DNA  
A:Residues: 1-358 <ST2>  
A:Cross-references: EMBL:X61352; NID:960186; PIDN:CAA3616.1; PID:960187  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer  
A:Reference number: S70417; MUID:92144209; PMID:11736940  
A:Accession: S70418  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333, 'X', 335-358 <ST2>  
A:Cross-references: EMBL:X61352; NID:960186  
C:Superfamily: type E retrovirus env polyprotein

Query Match 83.6%; Score 244; DB 2; Length 358;  
Best Local Similarity 80.8%; Pred. No. 8.2e-19;  
Matches 42; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWFWNI 52  
Db 126 DNMTWMEWDERIHNYTSLIYTLIESONQOEKNEDELLELDKMSLWFWNI 177

RESULT 15  
S22000

envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 12-Apr-1995  
C:Accession: S22000  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as dete

A:Reference number: S21990  
A:Accession: S22000  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <ST2>  
A:Cross-references: EMBL:X61351  
C:Superfamily: type E retrovirus env polyprotein

Query Match 83.6%; Score 244; DB 2; Length 358;  
Best Local Similarity 80.8%; Pred. No. 8.2e-19;  
Matches 42; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEDELLELDKMSLWFWNI 52  
Db 126 DNMTWMEWDERIHNYTSLIYTLIESONQOEKNEDELLELDKMSLWFWNI 177

Search completed: March 18, 2003, 14:37:17  
Job time : 17.6502 secs

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:27:26 ; Search time 7.42857 Seconds  
(without alignments)  
290.334 Million cell updates/sec

Title: US-09-877-606-3  
Perfect score: 292  
Sequence: 1 NHTTWLEWDREINNTSLIH.....NEQELLELDKASLMMWENI 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues  
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	99.3	856	1	ENV_HV1H2
2	290	99.3	856	1	ENV_HV1H3
3	290	99.3	856	1	ENV_HV1LW
4	277	94.9	851	1	ENV_HV1B8
5	277	94.9	856	1	ENV_HV1B1
6	277	94.9	861	1	ENV_HV1B8
7	274	93.8	853	1	ENV_HV1MF
8	268	91.8	856	1	ENV_HV1PV
9	254	87.0	856	1	ENV_HV1W1
10	253	86.6	856	1	ENV_HV1SC
11	252	86.3	852	1	ENV_HV1S3
12	251	86.0	847	1	ENV_HV1S1
13	251	83.2	847	1	ENV_HV1W2
14	242	82.9	856	1	ENV_HV1MN
15	241	82.5	853	1	ENV_HV1Z6
16	241	82.5	855	1	ENV_HV1Z2
17	241	82.5	867	1	ENV_HV1J3
18	240	82.2	865	1	ENV_HV1RH
19	238	81.5	852	1	ENV_HV1BN
20	237	81.2	853	1	ENV_HV1EL
21	236	80.8	855	1	ENV_HV1A2
22	236	80.8	861	1	ENV_HV1KB
23	234	80.1	846	1	ENV_HV1ND
24	233	79.8	868	1	ENV_HV1C4
25	232	79.5	843	1	ENV_HV1Y2
26	232	79.5	848	1	ENV_HV1JR
27	232	79.5	855	1	ENV_HV1OY
28	230	78.8	853	1	ENV_HV1Z8
29	229	78.4	859	1	ENV_HV1MA
30	213	72.9	856	1	ENV_HV1ZH
31	198	67.8	854	1	ENV_ST1VC2
32	142	48.6	854	1	ENV_ST1VC
33	130	44.5	882	1	ENV_ST1VM1

34	130	44.5	885	1	ENV_ST1VS4	P12492 simian immu
35	129	44.2	881	1	ENV_ST1VK	P05884 simian immu
36	127	43.5	889	1	ENV_ST1SP	P19503 simian immu
37	125	42.8	859	1	ENV_HV2ST	P20872 human immu
38	125	42.8	877	1	ENV_ST1VAG	P27977 simian immu
39	125	42.8	880	1	ENV_ST1VNL	P11267 simian immu
40	124	42.5	712	1	ENV_HV2S2	P32536 human immu
41	124	42.5	768	1	ENV_ST1V1	P27757 simian immu
42	123	42.1	859	1	ENV_HV2D2	P15831 human immu
43	120	41.1	859	1	ENV_HV2CA	P24105 human immu
44	120	41.1	860	1	ENV_HV2BE	P18094 human immu
45	119	40.8	851	1	ENV_HV2G1	P18040 human immu

## ALIGNMENTS

RESULT 1	ENV_HV1H2	STANDARD:	PRT:	856 AA.
ID	ENV_HV1H2			
AC	P04578; 009779;			
DT	13-AUG-1987 (Rel. 05, Created)			
DT	15-JUL-1999 (Rel. 38, Last sequence update)			
DT	30-MAY-2000 (Rel. 39, Last annotation update)			
DE	Envelope glycoprotein gp160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11706;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=87291916; PubMed=3040055;			
RA	Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,			
RA	Gallo R.C., Wong-Staal F.;			
RT	"Complete nucleotide sequences of functional clones of the AIDS virus."			
RL	AIDS Res. Hum. Retroviruses 3:57-69(1987).			
RN	(2)			
RP	REVISIONS.			
RA	Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,			
RA	Gallo R.C., Wong-Staal F.;			
RL	Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	EMBL: K03455; AAB50262.1; -			
DR	EMBL: AF038399; AAB99976.1; -			
DR	EMBL: AF033819; AAC82596.1; -			
DR	HIV: K03455; ENVSHXB2.			
DR	InterPro: IPR000328; Env.GP41.			
DR	InterPro: IPR000777; GP120.			
DR	Pfam: PF00516; GP120.1.			
KW	Pfam: PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; glycoprotein; Transmembrane; signal.			
FT	SIGNAL	1	30	
FT	CHAIN	31	511	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	DISULFID	512	856	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74	BY SIMILARITY.
FT	DISULFID	119	205	BY SIMILARITY.
FT	DISULFID	126	196	BY SIMILARITY.
FT	DISULFID	131	157	BY SIMILARITY.
FT	DISULFID	218	247	BY SIMILARITY.
FT	DISULFID	228	239	BY SIMILARITY.
FT	DISULFID	296	331	BY SIMILARITY.
FT	DISULFID	378	445	BY SIMILARITY.

FT DISULFID 385 418 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

Query Match 99.3%; Score 290; DB 1; Length 856;  
 Best Local Similarity 98.1%; Pred. No. 2.3e-24;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLEESNOQKNEOLLELDKMSLWMEFNI 52  
 DB 624 NHTTWLEMDREINNTSLIHSLEESNOQKNEOLLELDKMSLWMEFNI 675

RESULT 2  
 ENV\_HV1H3 STANDARD; PRT; 856 AA.  
 AC P04624;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide gp160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11707;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85228248; Pubmed=2988795;  
 RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R., Shaw G.M., Wong-Staal F., Reddy E.P.;  
 RT "HIV-III env gene products synthesized in E. coli are recognized by antibodies present in the sera of AIDS patients.";  
 RL Cell 41:979-986(1985).  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: M14100; AAA44679.1; -.  
 DR HIV: M14100; ENVSHXB3.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1  
 FT CHAIN 31  
 FT CHAIN 511  
 FT DISULFID 512 856 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT DISULFID 119 205 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 126 196 BY SIMILARITY.  
 FT DISULFID 131 157 BY SIMILARITY.  
 FT DISULFID 136 157 BY SIMILARITY.  
 FT DISULFID 181 247 BY SIMILARITY.  
 FT DISULFID 228 239 BY SIMILARITY.  
 FT DISULFID 296 331 BY SIMILARITY.  
 FT DISULFID 378 445 BY SIMILARITY.  
 FT DISULFID 385 418 BY SIMILARITY.  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 856 AA; 97188 MW; 3373C68B84C1AFC CRC64;

Query Match 99.3%; Score 290; DB 1; Length 856;  
 Best Local Similarity 98.1%; Pred. No. 2.3e-24;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLEESNOQKNEOLLELDKMSLWMEFNI 52  
 DB 624 NHTTWLEMDREINNTSLIHSLEESNOQKNEOLLELDKMSLWMEFNI 675

RESULT 3  
 ENV\_HV1LW STANDARD; PRT; 856 AA.  
 AC O70626;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide gp160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (WM12.3 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Retroviridae; Lentivirus.  
 CC -----

**us-09-877-606-3.rsp**

	Matches	51;	Conservative	1;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	NHTTTLLEMDREINNTTSLIHSILIESQNOQKNEBELLELDKMSLMMWPNFNI	52							
Db	624	NHTTMMEDREINNTTSLIHSILIESQNOQKNEBELLELDKMSLMMWPNFI	675							

ID	STANDARD;	PRT;	851 AA.
ENV-HV1BR			
P04582;			
AC			
DT 13-AUG-1987	(Rel. 05, Created)		
DT 13-AUG-1987	(Rel. 05, Last sequence update)		
DT 15-JUL-1999	(Rel. 38, Last annotation update)		
DE Envelope pol	protein gp160 precursor [Contains: Exterior membrane		
DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)]			
ENV			

OS Human immunodeficiency virus type 1 (Bt8 isolate) (HIV-1).  
OC Viruses; Retroviral viruses; Retroviridae.  
OX NCBI\_TaxID=11684;  
RN [1]  
RZ  
RX  
RP MEDLINE=85111123; PubMed=2578615;  
RA Ratner L., Haseltine W., Pataca R.J., Starcich B.R.,  
RJ Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,

RA Baumeister K., Ivanoff L., Pelletway S.R. Jr., Pearson M.L.,  
RA Lautenberger J. A., Papas T.S., Graybe V., Chang N.T., Gallo R.C.,  
RA Wong-Staal F.:  
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
RL Nature 313:277-284(1985).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC

CC The European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; K02011; AAA44661.1; -  
DR HIV; K02011; ENVSBB8.  
RR

DR GlycosylatedB: P04583; -,  
DR InterPro: IPR000328; Env\_GP41.  
DR InterPro: IPR00777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
DR AIDS: Coat protein; Polyprotein; Glycoprotein; Transmembrane;  
KW signal. 1 30  
TT SIGNAL

FT	CHAIN	31	506	EXTRIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	507	851	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54	74	BY SIMILARITY.
FT	DISULFID	119	205	BY SIMILARITY.
FT	DISULFID	126	196	BY SIMILARITY.
FT	DISULFID	131	157	BY SIMILARITY.
FT	DISULFID	218	247	BY SIMILARITY.
FT	DISULFID	228	239	BY SIMILARITY.

FT	DISULFID	296	331	BY SIMILARITY.	
FT	DISULFID	378	440	BY SIMILARITY.	
FT	DISULFID	385	413	BY SIMILARITY.	
FT	CARBOHYD	88		N-LINKED (GLCNAC.	.) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC.	.) (POTENTIAL).
FT	CARBOHYD	141	141	N-LINKED (GLCNAC.	.) (POTENTIAL).
FT	CARBOHYD	156	156	N-LINKED (GLCNAC.	.) (POTENTIAL).
FT	CARBOHYD	160	160	N-LINKED (GLCNAC.	.) (POTENTIAL).

FT	CARBOHYD	186	N-LINKED	(GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	197	N-LINKED	(GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	230	N-LINKED	(GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	234	N-LINKED	(GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	241	N-LINKED	(GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	241	N-LINKED	(GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	262	N-LINKED	(GLCNAC. . .)	(POTENTIAL)
FT	CARBOHYD	276	N-LINKED	(GLCNAC. . .)	(POTENTIAL)

```

FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MM; D16A3C90857785F1 CRC64;

Query Match 94.9%; Score 277; DB 1; Length 851;
Best Local Similarity 94.2%; Pred. No. 6.2e-23;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLHSLSIESONQOEKNEDELIDKMSLWMEFI 52
Db 619 NNMWTMEWDREINNTSLHSLSIESONQOEKNEDELIDKMSLWMEFI 670

RESULT 5
ENV_HV1B1 STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BR10 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentiviridae.
OX NCBI_Taxid=11678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Ratafski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R., Jr., Pearson M.L.,
RA Lutenberger J.A., Papas T.S., Chrayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).
RN [2]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
RT potential glycosylation sites of the type 1 recombinant human
RT immunodeficiency virus envelope glycoprotein (gp120) expressed in
RT Chinese hamster ovary cells";
RL J. Biol. Chem. 265:10373-10382(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M15654; AAAA4205.1; -
CC PIR: A03973; VCLJH3.
CC HIV: M15654; ENV5BH102.
CC InterPro: IPR000328; Env_Gp41.

```

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97224 MM; DBFPIA18931BB27 CRC64;

Query Match 94.9%; Score 277; DB 1; Length 856;
Best Local Similarity 94.2%; Pred. No. 6.2e-23;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLHSLSIESONQOEKNEDELIDKMSLWMEFI 52
Db 624 NNMWTMEWDREINNTSLHSLSIESONQOEKNEDELIDKMSLWMEFI 675

RESULT 6
ENV_HV1B1 STANDARD; PRT; 861 AA.
AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BR1 isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentiviridae.
OX NCBI_Taxid=11678;
RN [1]

```



[illegible]

QY	1	HTTTLTMMRELTNNYTSLIHSLIESONQOEKNEDELTLELDMWASIMMNT	52
DB	629	NNMTWMDRELTNNYTSLIHSLIESONQOEKNEDELTLELDMWASIMMNT	660
RESULT 7			
ENV_HV1MF			
ID	ENV_HV1MF	STANDARD:	PRT: 853 AA.
AC	P19551:		
DT	01-FEB-1991 (Rel. 17, Created)		
DT	01-FEB-1991 (Rel. 17, Last sequence update)		
DT	15-JUL-1999 (Rel. 38, Last annotation update)		
DE	Envelope polypeptide Gp160 precursor [Contains: Exterior membrane glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].		
CN	ENV.		
OS	Human immunodeficiency virus type 1 (MFA isolate) (HIV-1).		
OC	Viruses; Retroid viruses; Retroviridae; Lentiviruses.		
OX	NCBI_TaxID=11704;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=90317877; PubMed=1695254;		
RA	Stevenson M., Haggerty S., Lamontea C., Mann A.M., Meier C., Waslak A.;		
RA	"Cloning and characterization of human immunodeficiency virus type 1 variants diminished in the ability to induce syncytium-independent cytolysis.";		
RL	J. Virol. 64:3792-3803(1990).		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-stb.ch/announce/">http://www.isb-stb.ch/announce/</a> or send an email to <a href="mailto:license@isb-stb.ch">license@isb-stb.ch</a> ).		
CC	-----		
DR	EMBL, M33943; AAA44850.1; -.		
DR	HIV, M33943; ENVSMFA.		
DR	InterPro: IPR000328; Env_Gp41.		
DR	InterPro: IPR000777; Gp120.		
DR	Pfam: PF00516; Gp120.1.		
KM	Pfam: PF00517; Gp41.1.		
KM	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; signal.		
KM	SIGNAL.		
FT	CHAIN	1 30	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	CHAIN	510 853	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	54 74	BY SIMILARITY.
FT	DISULFID	119 203	BY SIMILARITY.
FT	DISULFID	126 194	BY SIMILARITY.
FT	DISULFID	131 157	BY SIMILARITY.
FT	DISULFID	216 245	BY SIMILARITY.
FT	DISULFID	226 237	BY SIMILARITY.
FT	DISULFID	294 329	BY SIMILARITY.
FT	DISULFID	376 443	BY SIMILARITY.
FT	DISULFID	383 416	BY SIMILARITY.
FT	CARBOHYD	88 88	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	136 136	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	141 141	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	156 156	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	160 160	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	186 186	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	195 195	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	232 232	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	239 239	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	260 260	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	274 274	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	287 287	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	293 293	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	329 330	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	330 330	N-LINKED (GLCNAc . . .) (POTENTIAL).
FT	CARBOHYD	354 354	N-LINKED (GLCNAc . . .) (POTENTIAL).

FT	DISULFID	296	331	BY SIMILARITY.
FT	DISULFID	378	445	BY SIMILARITY.
FT	DISULFID	385	418	BY SIMILARITY.
FT	CARBOHYD	88	88	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	141	141	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	156	156	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	160	160	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	186	186	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	197	197	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	230	230	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	234	234	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	241	241	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	262	262	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	276	276	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	289	289	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	295	295	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	301	301	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	339	339	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	386	386	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	392	392	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	397	397	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	406	406	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	448	448	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	463	463	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	611	611	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	625	625	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	637	637	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	674	674	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	750	750	N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD	816	816	N-LINKED (GLCNAC. . . ) (POTENTIAL).
SO	SEQUENCE	856 AA;	97339 MW;	5FCDB1DC31C2109B3 CRC64;
Query Match		91.8%;	Score 268;	DB 1; Length 856;
Best Local Similarity		90.4%;	Pred. No. 6,1e-22;	
Matches 47;	Conservative	3;	Mismatches 2;	Indels 0; Gaps 0;
QY	1	NHTTWLEWDREINNYTSLHSLIESONQOEKNEDELLEDKMSLWMMFNI	52	
Db	624	NNMTWMEWDREINNYTSLHSLIESONQOEKNEDELLEDKMNLWMLNI	675	
RESULT 9				
ENV_HV1W1				
ID	ENV_HV1W1	STANDARD;	PRT:	856 AA.
AC	P31872;			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	01-JUL-1993 (Rel. 26, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Envelope polypeptide GP160 precursor [contaminant: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
DE	glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (WMJ1 isolate) (HIV-1).			
OC	Viruses; Retroviral viruses; Retroviridae; Lentivirus.			
OX	NCBI_Taxid=31678;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=66218077; PubMed=2423250;			
RA	Starck B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H., Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;			
RT	"Identification and characterization of conserved and variable regions in the envelope gene of HTLV-III/LAV, the retrovirus of			
RT	AIDS."			
RL	Cell 45:637-648(1986).			
CC	-1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO WAS PERINATALLY INFECTED BY HER MOTHER.			
CC	PIR: A24774; VCLJ3W.			
DR	Interpro: IPR000328; Env_GP41.			
DR	Interpro: IPR000777; GP120.			

[illegible][illegible]



```

FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 434 434 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 847 AA; 96135 MW; 0A901317FD7FF2AB CRC64;

```

```

Query Match 86.0%; Score 251; DB 1; Length 847;
Best Local Similarity 82.7%; Pred. No. 4.5e-20;
Matches 43; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 NHTTWLEMDREINNTSLHSLEESONQOEKNEDELLELDKWSLMMWNI 52
Db 615 NNTTWMEEREIDNTNLTLYLIEESONQOEKNEDELLELDKWSLMMWNI 666

```

```

RESULT 13
ID ENV_HV1W2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks W.P.;
RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or
at risk for AIDS."
RL Science 232:1548-1553(1986).
CC -1 MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC WAS PERMANENTLY INFECTED BY HER MOTHER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, M12507; AAB12990.1; -
CC HIV; M12507; ENV5WMJ2.
CC InterPro: IPR000328; Env.GP41.
CC InterPro: IPR000777; GP120.
CC Pfam: PF00516; GP120; 1.
CC Pfam: PF00517; GP41; 1.
CC AIDS; Coat protein; Glycoprotein; glycoprotein; Transmembrane;
CC signal.

```

```

FT SIGNAL 1 29
FT CHAIN 30 501
FT CDR1 502 847
FT DISULFID 553 773
FT DISULFID 118 202
FT DISULFID 125 193
FT DISULFID 130 152
FT DISULFID 215 244
FT DISULFID 225 326
FT DISULFID 293 326
FT DISULFID 372 435
FT DISULFID 379 408
FT DISULFID 87 87
FT CARBOHYD 134 134
FT CARBOHYD 140 140
FT CARBOHYD 151 151
FT CARBOHYD 155 155
FT CARBOHYD 183 183
FT CARBOHYD 184 184
FT CARBOHYD 194 194
FT CARBOHYD 231 231
FT CARBOHYD 238 238
FT CARBOHYD 259 259
FT CARBOHYD 273 273
FT CARBOHYD 286 286
FT CARBOHYD 292 292
FT CARBOHYD 327 327
FT CARBOHYD 334 334
FT CARBOHYD 350 350
FT CARBOHYD 356 356
FT CARBOHYD 380 380
FT CARBOHYD 386 386
FT CARBOHYD 390 390
FT CARBOHYD 400 400
FT CARBOHYD 438 438
FT CARBOHYD 450 450
FT CARBOHYD 602 602
FT CARBOHYD 607 607
FT CARBOHYD 616 616
FT CARBOHYD 628 628
SQ SEQUENCE 847 AA; 96466 MW; CD1E3D73A53BCAE CRC64;

```

```

Query Match 83.2%; Score 243; DB 1; Length 847;
Best Local Similarity 80.8%; Pred. No. 3.4e-19;
Matches 42; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

```

```

OY 1 NHTTWLEMDREINNTSLHSLEESONQOEKNEDELLELDKWSLMMWNI 52
Db 615 DNTTWMEEREIDNTNLTLYLIEESONQOEKNEDELLELDKWSLMMWNI 666

```

```

RESULT 14
ID ENV_HV1WN STANDARD; PRT; 856 AA.
AC P05877;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11696;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88219542; PubMed=3369091;
RA Gurgo C., Guo H.-G., Franchini G., Aldovini A., Collalti E.,
RA Farrell K., Wong-Staal F., Gallo R.C., Reitz M.S. Jr.;
RT "Envelope sequences of two new United States HIV-1 isolates."
RL Virology 164:531-536(1988).
CC -1 MISCELLANEOUS: THE MN ISOLATE WAS TAKEN FROM A PEDIATRIC AIDS
CC PATIENT IN 1984.

```



FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 853 AA; 97043 MW; 849B08BCBAFF7008 CRC64;

Query Match  
 Best Local Similarity 82.5%; Score 241; DB 1; Length 853;  
 Matches 42; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 4 TWLEWDREINNYTSLIHSLEESQNOQEKNEQELLELDKWSLNNWFNI 52  
 ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:  
 Db 624 TMMWEEREIDNYTGILYRLIEESQTOQEKNEQELLELDKWSLNNWFNI 672

Search completed: March 18, 2003, 14:34:01  
 Job time : 8.42857 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:01 ; Search time 29.202 Seconds  
(without alignments)  
366.908 Million cell updates/sec

Title: US-09-877-606-3  
Perfect score: 292  
Sequence: 1 NHTTWLEWDREINNTSLIH.....NEQELLEDKNASIMNMFNI 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*

- 1: sp.archaea:\*
- 2: sp.bacteria:\*
- 3: sp.fungi:\*
- 4: sp.human:\*
- 5: sp.invertebrate:\*
- 6: sp.mammal:\*
- 7: sp.mhc:\*
- 8: sp.organelle:\*
- 9: sp.phage:\*
- 10: sp.plant:\*
- 11: sp.todent:\*
- 12: sp.virus:\*
- 13: sp.vertebrate:\*
- 14: sp.unclassified:\*
- 15: sp.virus:\*
- 16: sp.bacteriap:\*
- 17: sp.archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	99.3	748	15	Q70606 human immun
2	290	99.3	752	15	Q70604 human immun
3	290	99.3	752	15	Q70605 human immun
4	286	97.9	747	15	Q70607 human immun
5	286	97.9	752	15	Q70608 human immun
6	285	97.6	856	15	Q90SM7 human immun
7	282	96.6	852	15	Q89797 human immun
8	279	95.5	856	15	Q92877 human immun
9	277	94.9	645	15	Q993A6 human immun
10	277	94.9	851	15	Q782A3 human immun
11	277	94.9	854	15	Q85582 human immun
12	277	94.9	854	15	Q72502 human immun
13	277	94.9	856	15	Q74589 human immun
14	273	93.5	856	15	Q74090 human immun
15	273	93.5	854	15	Q90178 human immun
16	273	93.5	854	15	Q78705 human immun

17	266	91.1	757	15	Q9Q722 human immun
18	265	90.8	855	15	Q9E1R7 human immun
19	264	90.4	848	15	Q69990 human immun
20	263	90.1	851	15	Q56110 human immun
21	262	89.7	727	15	Q9Q723 human immun
22	261	89.4	616	15	Q993B0 human immun
23	261	89.4	618	15	Q993B2 human immun
24	260	89.0	358	15	Q78120 human immun
25	260	89.0	848	15	Q69988 human immun
26	259	88.7	635	15	Q90U82 human immun
27	259	88.7	838	15	Q03806 human immun
28	259	88.7	852	15	Q69992 human immun
29	259	88.7	854	15	Q78225 human immun
30	259	88.7	855	15	Q03805 human immun
31	258	88.4	862	15	Q9E1S2 human immun
32	257	88.0	856	15	Q72993 human immun
33	257	88.0	859	15	Q8Q863 human immun
34	256	87.7	851	15	Q8Q852 human immun
35	256	87.7	863	15	Q9WJ04 human immun
36	256	87.7	864	15	Q9YP48 human immun
37	255	87.3	847	15	Q69996 human immun
38	255	87.3	854	15	Q56112 human immun
39	255	87.3	863	15	Q9WJ08 human immun
40	255	87.3	863	15	Q42031 human immun
41	255	87.3	864	15	Q9WJ01 human immun
42	254	87.0	849	15	Q77368 human immun
43	253	86.6	884	15	Q91K06 human immun
44	253	86.6	847	15	Q41537 human immun
45	253	86.6	855	15	Q91K00 human immun

#### ALIGNMENTS

RESULT 1

Q70606 PRELIMINARY: PRT: 748 AA.

AC Q70606: 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)

DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DT Envelope glycoprotein (Fragment).

ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_TaxID=11676;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=LM881;

RX MEDLINE=95127297; PubMed=7826699;

RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B., Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W., "Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HTLV type IIIB).";

RT AIDS Res. Hum. Retroviruses 10:1143-1155(1994).

RL [2]

RN SEQUENCE FROM N.A.

RP STRAIN=LM881;

RC Mulder K.E.;

RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.

DR EMBL: U12032; AAA76668.1; -

DR InterPro: IPR000328; Env-GP41.

DR InterPro: IPR00777; GP120.

DR Pfam: PF00516; GP120.1.

DR Pfam: PF00517; GP41.1.

KW AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.

FT NON-TER

SQ SEQUENCE 748 AA; 84224 MW; 56BEDF186C67694B CRC64;

Query Match 99.3%; Score 290; DB 15; Length 748;  
Best Local Similarity 98.1%; Pred. No. 1.6e-22;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 NHTTWLEMDREINNTSLIHSIESQNOQKNEQELLELDKWSLWMWFI 52
    |||||
Db 620 NHTTWEMDREINNTSLIHSIESQNOQKNEQELLELDKWSLWMWFI 671

RESULT 2
OY 070604 PRELIMINARY; PRT; 752 AA.
AC 070604;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LM851;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
  Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
  infected with HIV type 1 (HTLV type IIIB).";
RN [2]
RP AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LM852;

```

```

RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12031; AAA76667.1; -.
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
FT NON_TER 752
SO SEQUENCE 752 AA; 84894 MM; 8B30AE894013B45A CRC64;

Query Match 99.3%; Score 290; DB 15; Length 752;
Best Local Similarity 98.1%; Pred. No. 1.6e-22;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSIESQNOQKNEQELLELDKWSLWMWFI 52
    |||||
Db 624 NHTTWEMDREINNTSLIHSIESQNOQKNEQELLELDKWSLWMWFI 675

RESULT 4
OY 070607 PRELIMINARY; PRT; 747 AA.
AC 070607;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LM87-1;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
  Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
  infected with HIV type 1 (HTLV type IIIB).";
RN [2]
RP AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LM87-1;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL; U12034; AAA76669.1; -.
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.
FT NON_TER 747
SO SEQUENCE 747 AA; 84250 MM; 732E836A52245F14 CRC64;

Query Match 97.9%; Score 286; DB 15; Length 747;
Best Local Similarity 98.1%; Pred. No. 4.2e-22;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSIESQNOQKNEQELLELDKWSLWMWFI 52
    |||||
Db 619 NHTTWEMDREINNTSLIHSIESQNOQKNEQELLELDKWSLWMWFI 670

RESULT 5
OY 070608 PRELIMINARY; PRT; 752 AA.
AC 070608;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Envelope glycoprotein (Fragment).
GN ENV.

```

```

OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LM87-2;
RA MEDLINE=95127297; PubMed-7826699;
RX Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
RT "Viral variability and serum antibody response in a laboratory worker
infectd with HIV type 1 (HIV type IIIB).";
RN AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LM87-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U12053; AAA7670.1; -
DR InterPro: IPR000328; Env_Gp41.
DR InterPro: IPR000777; Gp120.
DR Pfam: PR00516; Gp120.1.
DR Pfam: PR00517; Gp41.1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER
SQ SEQUENCE 752 AA; 84780 MW; 708672A2DC0CE8F8 CRC64;

Query Match          97.9%; Score 286; DB 15; Length 752;
Best Local Similarity 98.1%; Pred. No. 4.2e-22;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWEMREINNTYSLHSLIEESONOEKNEDELLEDKWASIMNPNFI 52
Db 624 NHTTWEMREINNTYSLHSLIEESONOEKNEDELLEDKWASIMNPNFI 675

RESULT 6
ID 090SM7 PRELIMINARY; PRT; 856 AA.
AC 090SM7;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HXB2;
RA Altman-Onal Y., Cheynet V., Verrier B.;
RT "Mutations and transcriptional alterations associated with the
downregulation of HIV-1 envelope glycoprotein expression following
acute cytopathic effects.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF358141; AAK49977.1; -
DR InterPro: IPR000328; Env_Gp41.
DR InterPro: IPR000777; Gp120.
DR Pfam: PR00516; Gp120.1.
DR Pfam: PR00517; Gp41.1.
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 856 AA; 97126 MW; 9458D02B2FD734B3 CRC64;

Query Match          97.6%; Score 285; DB 15; Length 856;
Best Local Similarity 96.2%; Pred. No. 6.1e-22;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWEMREINNTYSLHSLIEESONOEKNEDELLEDKWASIMNPNFI 52
Db 624 NHTTWEMREINNTYSLHSLIEESONOEKNEDELLEDKWASIMNPNFI 675

RESULT 7

```

```

089797
ID 089797 PRELIMINARY; PRT; 852 AA.
AC 089797;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LM90-2;
RX MEDLINE=95127297; PubMed-7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
RA Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
infectd with HIV type 1 (HIV type IIIB).";
RN AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LM90-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U12053; AAA76685.1; -
DR InterPro: IPR000328; Env_Gp41.
DR InterPro: IPR000777; Gp120.
DR Pfam: PR00516; Gp120.1.
DR Pfam: PR00517; Gp41.1.
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;
Transmembrane.
SQ SEQUENCE 852 AA; 96347 MW; 4E865229D4EB33CF CRC64;

Query Match          96.6%; Score 282; DB 15; Length 852;
Best Local Similarity 96.2%; Pred. No. 1.3e-21;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWEMREINNTYSLHSLIEESONOEKNEDELLEDKWASIMNPNFI 52
Db 620 NHTTWEMREINNTYSLHSLIEESONOEKNEDELLEDKWASIMNPNFI 671

RESULT 8
ID 092877 PRELIMINARY; PRT; 856 AA.
AC 092877;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Simian-Human immunodeficiency virus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=57667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9908984; PubMed-9882298;
RA Cayabyab M., Karlsson G.B., Etamad-Moghadam B.A., Hofmann W.,
RA Steenbeke T., Halloran M., Fanton J.W., Axthelm M.K., Letvin N.L.,
RA Sodroski J.G.;
RT "Changes in human immunodeficiency virus type 1 envelope glycoproteins
responsible for the pathogenicity of a multiply passaged simian-human
immunodeficiency virus (SHV-HXBc2).";
RN J. Virol. 73:976-984(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Cayabyab M., Karlsson G.B., Etamad-Moghadam B., Hofmann W.,
RA Halloran M., Axthelm M.W., Letvin N.L., Sodroski J.G.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF041850; AAD12142.1; -
DR InterPro: IPR000328; Env_Gp41.

```

DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Transmembrane.  
SQ SEQUENCE 856 AA; 97151 MM; C50BE0388FB73659 CRC64;  
Query Match 95.5%; Score 279; DB 15; Length 856;  
Best Local Similarity 94.2%; Pred. No. 2.6e-21;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 1 NHTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFNI 52  
DB 624 NHTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFNI 675  
RESULT 9  
O993A6 PRELIMINARY: PRT: 645 AA.  
AC O993A6.  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Truncated envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1007;  
RX MEDLINE=21192672; PubMed=11287644;  
RA Suman S., Lockey T.D., Stobod K.S., Jones B., Riberdy J.M.,  
White S.W., Doherty P.C., Hurwitz J.L.;  
RT "Localization of CD4+ T cell epitope hotspots to exposed strands of  
RT HIV envelope glycoprotein suggests structural influences on antigen  
RT processing";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4587-4592(2001).  
DR EMBL: AF321563; AAK18810.1; -.  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KW AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
FT NON\_TER 1  
SQ SEQUENCE 645 AA; 72485 MM; B076514BE93362EC CRC64;  
Query Match 94.9%; Score 277; DB 15; Length 645;  
Best Local Similarity 94.2%; Pred. No. 3.2e-21;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 1 NHTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFNI 52  
DB 593 NHTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFNI 644  
RESULT 10  
O782A3 PRELIMINARY: PRT: 851 AA.  
AC O782A3.  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Env polyprotein.  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Carlini F., Federici M., Equestre M., Ricci S., Ratti G., Zibai Q.,  
RA Verani P., Rossi G.B.;  
RT "Sequence analysis of HIV-1 proviral DNA from a non producer

RT Chronically infected HUT-78 cellular clone.";  
RL J. Viral Diseases 1:40-55(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89352106; PubMed=2765297;  
RA Federico M., Titti F., Butto S., Orecchia A., Carlini F., Taddeo B.,  
Maccia B., Mangano N., Verani P., Rossi G.;  
RT "Biological and molecular characterization of producer and non  
RT producer clones from HUT-78 infected with a patient HIV isolate.";  
RL AIDS Res. Hum. Retroviruses 5:385-396(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Titti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,  
Borsetti A., Saglio I., Verani P., Rossi G.;  
RT "Variability of HIV-1 virus: characteristics of an infected but not  
RT productive clone.";  
RL Int. J. Immunopharmacol. 3:17-23(1990).  
DR EMBL: Z11530; CAA77628.1; -.  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 851 AA; 96630 MM; 1A3767B5B7E98027 CRC64;  
Query Match 94.9%; Score 277; DB 15; Length 851;  
Best Local Similarity 94.2%; Pred. No. 4.3e-21;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
OY 1 NHTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFNI 52  
DB 619 NHTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFNI 670  
RESULT 11  
O85582 PRELIMINARY: PRT: 854 AA.  
AC O85582.  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Envelope polyprotein.  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86281827; PubMed=3016298;  
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,  
Martin M.A.;  
RT "Production of acquired immunodeficiency syndrome-associated  
RT retrovirus in human and nonhuman cells transfected with an infectious  
RT molecular clone.";  
RL J. Virol. 59:284-291(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;  
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Buckler C.E.;  
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92219406; PubMed=1373204;  
RA Dai L.C., Litaua R., Takahashi K., Ennis F.A.;  
RT "Mutation of human immunodeficiency virus type 1 at amino acid 585 on  
RT gp41 results in loss of killing by CD8+ A24-restricted cytotoxic T  
RT lymphocytes.";  
RL J. Virol. 66:3151-3154(1992).  
DR EMBL: M19921; AAA44992.1; -.  
DR InterPro: IPR000328; Env\_GP41.

DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.  
 SO SEQUENCE 854 AA; 97124 MW; ABA0A1BF36D5595A CRC64;

Query Match Best Local Similarity 94.9%; Score 277; DB 15; Length 854;  
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLHSLIEESONQOEKNEQELLELDKASLMMWFI 52  
 Db 622 NNTWMEDEINNYTSLHSLIEESONQOEKNEQELLELDKASLMMWFI 673

## RESULT 12

OY 072502 PRELIMINARY; PRT; 854 AA.  
 AC 072502;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ENV polypotein.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovird viruses; Retrovirdae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MDA-3;  
 RX MEDLINE=96036482; PubMed=7483282;  
 RA Salminen M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,  
 RA Michael N.L., Carr J.K., Burke D.S., McCutchen F.E.;  
 RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes  
 RT from primary virus cultures using the polymerase chain reaction.";  
 RT virology 213:80-86(1995).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=MDA-3;  
 RX MEDLINE=86281827; PubMed=3016298;  
 RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,  
 RA Martin M.A.;  
 RT "Production of acquired immunodeficiency syndrome-associated  
 RT retrovirus in human and nonhuman cells transfected with an infectious  
 RT molecular clone.";  
 RT J. Virol. 59:284-291(1986).  
 RL EMBL: U26942; AAB60578.1;  
 DR InterPro: IPR000328; Env-GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.  
 FT CONFLICT 214 H -> L (IN REF. 2).  
 FT CONFLICT 530 A -> S (IN REF. 2).  
 FT CONFLICT 739 G -> D (IN REF. 2).  
 SO SEQUENCE 854 AA; 97005 MW; FF2264B3841D1220 CRC64;

Query Match Best Local Similarity 94.9%; Score 277; DB 15; Length 854;  
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLHSLIEESONQOEKNEQELLELDKASLMMWFI 52  
 Db 622 NNTWMEDEINNYTSLHSLIEESONQOEKNEQELLELDKASLMMWFI 673

## RESULT 13

ID 074599 PRELIMINARY; PRT; 856 AA.  
 AC 074599;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE ENV.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovird viruses; Retrovirdae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MDA-3;  
 RX MEDLINE=90101366; PubMed=1688473;  
 RA Cloyd M.W., Moore B.E.;  
 RT "Spectrum of Biological Properties of Human Immunodeficiency Virus  
 RT (HIV-1) Isolates.";  
 RT virology 174:103-116(1990).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=MDA-3;  
 RA Iwataki Y.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: D86068; BAA12995.1;  
 DR InterPro: IPR000328; Env-GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.  
 SO SEQUENCE 856 AA; 97287 MW; 238042A234C56685 CRC64;

Query Match Best Local Similarity 94.9%; Score 277; DB 15; Length 856;  
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLHSLIEESONQOEKNEQELLELDKASLMMWFI 52  
 Db 624 NNTWMEDEINNYTSLHSLIEESONQOEKNEQELLELDKASLMMWFI 675

## RESULT 14

ID 074090 PRELIMINARY; PRT; 856 AA.  
 AC 074090;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ENV.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retrovird viruses; Retrovirdae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MDA-3;  
 RX MEDLINE=90101366; PubMed=1688473;  
 RA Cloyd M.W., Moore B.E.;  
 RT "Spectrum of Biological Properties of Human Immunodeficiency Virus  
 RT (HIV-1) Isolates.";  
 RT virology 174:103-116(1990).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=MDA-3;  
 RA Iwataki Y.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: D86069; BAA13003.1;  
 DR InterPro: IPR000328; Env-GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.  
 SO SEQUENCE 856 AA; 97396 MW; FE3E784C423C108C CRC64;

Query Match Best Local Similarity 94.9%; Score 277; DB 15; Length 856;  
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLHSLIEESONQOEKNEQELLELDKASLMMWFI 52



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:52:44 ; Search time 1465.73 Seconds  
(without alignments) 574.569 Million cell updates/sec

Title: US-09-877-606-3

Perfect score: 292

Sequence: 1 NHTTWLEMDREINNTSLIH.....NQELLELDKMASLNNWFNI 52

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters: -DEV=tblh  
-MODE=frame+;p2n.model -DEV=tblh  
-Q=/cgn2.1/USFTO.spool/US09877606/runat\_18032003\_134800\_17705/app.query.fasta\_1.1059  
-DB=EST -QFMT=fastap -SUFIX=-rst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITs-bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human4.0.cdi -LIST=45  
-DOCAALIGN=200 -THR.SCORE=dot -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09877606 @CGN1\_1.1 4134 @runat\_18032003\_134800\_17705 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAPP -LARGEORDER -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMECUT=120  
-WARN\_TIMECUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estcin:\*  
4: em\_estcnu:\*  
5: em\_estcov:\*  
6: em\_estcpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_estcl:\*  
10: gb\_estl2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_frod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	71.5	24.5	491	17	AQ975040
C 2	70.5	24.1	498	17	A2554114
C 3	70	24.0	559	17	AQ778007
C 4	69	23.6	549	9	A1878626
C 5	69	23.6	1116	17	CNS06803
C 6	69	23.6	1133	17	CNS06823
C 7	68	23.3	345	14	BQ320438
C 8	67.5	23.1	221	13	BM531221
C 9	67.5	23.1	390	13	BM081133
C 10	67.5	23.1	404	13	BM140718
C 11	67.5	23.1	508	13	BJ180020
C 12	67.5	23.1	576	13	BJ063369
C 13	67.5	23.1	606	12	BC046100
C 14	67	22.9	454	13	BM574215
C 15	67	22.9	495	9	AA438382
C 16	67	22.9	590	12	BC384309
C 17	67	22.9	596	13	BI220845
C 18	67	22.9	610	17	AQ178922
C 19	67	22.9	842	17	BM424342
C 20	66.5	22.8	554	17	BM267779
C 21	66.5	22.8	569	14	BM882520
C 22	66.5	22.8	588	17	AQ710961
C 23	66.5	22.8	606	9	AU021710
C 24	66	22.6	394	9	A1944599
C 25	66	22.6	536	17	AQ437075
C 26	66	22.6	577	13	BM275249
C 27	66	22.6	601	17	BH280804
C 28	66	22.6	704	17	BH37421
C 29	66	22.6	749	12	BE784950
C 30	66	22.6	779	12	BG673357
C 31	65.5	22.4	248	9	AV068114
C 32	65.5	22.4	259	10	BB014574
C 33	65.5	22.4	398	13	BI672809
C 34	65.5	22.4	480	17	CNS00060
C 35	65.5	22.4	529	9	A1562605
C 36	65.5	22.4	708	9	AJ450148
C 37	65	22.3	373	12	BG137622
C 38	65	22.3	406	9	A1676261
C 39	65	22.3	444	12	BG136846
C 40	65	22.3	511	12	BG140409
C 41	65	22.3	531	14	BQ490224
C 42	65	22.3	785	17	BH669477
C 43	65	22.3	789	13	BM162888
C 44	65	22.3	995	17	CNS01619
C 45	64.5	22.1	381	14	D48829

# ALIGNMENTS

RESULT 1  
LOCUS AQ975040/c 491 bp DNA linear GSS 28-JAN-2000  
DEFINITION RPI-23-317P8.TJ RPI-23 Mus musculus genomic clone RPI-23-317P8,  
ACCESSION AQ975040  
VERSION AQ975040.1 GI:6805494  
KEYWORDS GSS.  
SOURCE house mouse.  
ORGANISM Mus musculus.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 491)  
Zhao,S., Nierman,W., Feldlyum,T., Malek,J., Shatsman,S., Aklnret,  
B., Levins,M., Moggann,S., Tsegaye,G., Geer,K., Kroll,M., de Jong,P.





VERSION	AQ778007.1	GI:5680967
KEYWORDS	GSS.	
SOURCE	human.	
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. 1 (bases 1 to 559)	
REFERENCE	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,W.D. and Hood,L.	
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)	
MEDLINE	99380589	
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones may be purchased from Research Genetics (info@resgen.com). BAC end Web Server: http://www.hscs.washington.edu Plate: 3119 row: C column: 5 Seq primer: T7 Class: BAC ends High quality sequence stop: 559. Location/Qualifiers 1..559 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="plate=3119 col=5 Row=C" /clone_lib="CIT Approved Human Genomic Sperm Library D" /sex="male" /note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in E-Coli DH10B"	
BASE COUNT	157 a 121 c 143 g 133 t	5 others
ORIGIN		
Alignment Scores:	Pred. NO.:	Length: 559
Score:	60.1	Matches: 20
Percent Similarity:	70.00	Conservative: 9
Best Local Similarity:	42.65%	Mismatches: 15
Query Match:	29.41%	Indels: 24
DB:	23.97%	Gaps: 3
US-09-877-606-3 (1-52) x AQ778007 (1-559)		
OY	2 histrrtTTPLeugluTFrPspArgcluiIeasnAsrtYThSer-----	17
Db	:   ::    ::   ::    :	
OY	18 -----	-leuileHis 20
Db	343 GCATGATGTTCCTCGTAGTGAGAGTGGGGTACTACACANGACTGACCATTGTTCTAT	402
OY	21 SerLeuilegluglu-----SerGlnasnngIngluIuYasngluIngluleuLeu 38	
Db	::    ::	
OY	39 GlueuuasplystTpPalaserleu 46	
Db	463 CCTATCGAAACATGGAATATGTGTG 486	
RESULT 4		
LOCUS	AI878626	549 bp mRNA linear EST 07-JUN-2001
DEFINITION	fcg2ph07.x1 Zebrafish Washu MPMGc EST Danio rerio cdna clone	
ACCESSION	AI878626	
VERSION	AI878626.1	GI:5552675
KEYWORDS	EST.	

```

SOURCE          zebrafish.
ORGANISM        Danio rerio
Euarchyota: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
Cyprinidae; Danio.
REFERENCE       Clark,M., Johnson,S.L., Lehrach,H., Lee,R., Li,F., Marra,M., Eddy
AUTHORS         S., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood
                ,K., Stepien,M., Theising,B., Allen,M., Bowers,Y., Person,B.,
                Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk,R., Ritter,E.,
                Korn,S., Shih,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R.
                and Wilson,R.
TITLE           Washu zebrafish EST Project 1998
JOURNAL          Unpublished (1998)
COMMENT          Contact: Stephen L. Johnson
                Washington University School of Medicine
                4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
                Tel.: 314 286 1800
                Fax: 314 286 1810
                Email: zbtrafish@watson.wustl.edu
cdna Library Preparation: Matthew Clark. cDNA Library Arrayed by:
Matthew Clark. DNA Sequencing by: Washington University Genome
Sequencing Center Clone distribution: Genome Systems, St. Louis,
Missouri (web address: www.genomesystems.com) (email contact:
info@genomesystems.com) and Research Genetics, Huntsville, Alabama
(web address: www.resgen.com) (email contact: info@resgen.com) and
RessourcenZentrumImPrimateDenbank, Berlin, Germany (web address:
www.rzp.de)
Seq primer: 'T7 ET from Amershams'
High quality sequence stop: 429.

FEATURES
    source
        location/qualities
            1..549
                /organism="Danio rerio"
                /db_xref="taxon:7955"
                /clone="IMAGE:3726013"
                /clone_lib="zebrafish Washu MPMG EST"
                /sex="mixed"
                /tissue_type="26 somite embryos, adult livers, shield
                    stage embryos"
                /lab_host="XLI-blue MRP"
                /note="Vector: pSPORT1; Site.1: NotI; Site.2: SalI; 1st
                    strand cDNA was primed with a Not I - oligo(dT)15 primer
                    [5'-GCAGTAGTCCTAGATCGGACGGCCGCCCTTTTTTTTTT-3'];
                    double-stranded cDNA was ligated to Sal I adaptors (BRL),
                    digested with Not I and cloned into the Not I and Sal I
                    sites of the pSPORT1 vector (BRL). Library was constructed
                    by Matthew Clark (Lehrach lab; ICPE, London and Max Planck
                    Institut fuer Molekulare Genetik Berlin). cDNAs for EST
                    analysis were selected following oligonucleotide
                    hybridization fingerprinting of arrayed clones from
                    zebrafish late somitogenesis (26 ss), adult liver or
                    embryonic shield stage (5.6 h) libraries. Fingerprint
                    data were used to computationally cluster cDNAs, and a
                    single cDNA from each cluster was chosen for sequencing.
                    In some cases multiple members of the same cluster were
                    sequenced to assess clustering parameters or single clones
                    were sequenced additional times to assess quality
                    control."
BASE COUNT      213 a             102 c              83 g             151 t
ORIGIN
Alignment Scores:
Pred. NO.:      78.3               Length:                 549
Score:          69.00              Matches:                  14
Percent Similarity: 51.16%          Conservative:             8
Best Local Similarity: 32.56%       Mismatch penalties:     17
Query Match:    23.63%              Indels:                   4
DB:             9                      Gaps:                     1
US-09-877-606-3 (1-52) x A1878626 (1-549)
yy   5 TrrpleugltpparpargglulleksnAsnyfThrSerleullHlsSerleullEgu 24
||||| | ||||| | ||||| : ::::| | ||| : ::| :|

```

Db 108 TGGCTGACATGCTCAAAAAATCAACCCGTTTCAGTCTACTACAGACACATGACGAG 167

Qy 25 GluserglnAsngIngluL-----LysAsngIngluIngluLeuLeu 40  
 :: ::::: ||||| ::::: ||||| ::::: |||||

Db 168 AAAAGAGAAATCCCAAGAGCTCATTCATTAAATGAGAAACACACATACAGGCTC 227

Qy 41 AsplyStrp 43  
 |||

Db 228 ACATGCTGG 236

RESULT 5  
 CNS06ROC/c 1116 bp DNA linear GSS 05-JUL-2001  
 LOCUS T7 end of clone AM0A016F08 of library AM0A from strain CLIB 89 of  
 DEFINITION Yarrowia lipolytica, genomic survey sequence.  
 AL412282  
 AL412282.1 GI:12182678  
 GSS.  
 Yarrowia lipolytica.  
 Yarrowia lipolytica  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Dipodascaceae; Yarrowia.  
 Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
 1 (bases 1 to 1116)  
 Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,  
 de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,  
 Malpertuy,A., Neuvéglise,C., Ozler-Kalogeropoulos,O., Potier,S.,  
 Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,  
 Wincker,P. and Weissenbach,J.  
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
 yeast species for molecular evolution studies  
 FEBS Lett. 487 (1), 3-12 (2000)

JOURNAL  
 MEDLINE 20584711  
 PUBMED 11152876

REFERENCE  
 AUTHORS 2 (bases 1 to 1116)  
 Casaregola,S., Neuvéglise,C., Lepingle,A., Bon,E., Feynerol,C.,  
 Artiguenave,F., Wincker,P. and Gallardin,C.  
 Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia  
 lipolytica  
 FEBS Lett. 487 (1), 95-100 (2000)

JOURNAL  
 MEDLINE 20584727  
 PUBMED 11152892

REFERENCE  
 AUTHORS 3 (bases 1 to 1116)  
 Genoscope.  
 Direct Submission  
 Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,  
 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
 segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT  
 JOURNAL  
 TITLE This GSS is part of a random genomic sequencing program of thirteen  
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
 exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,  
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,  
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  
 5 kb were prepared and both extremities were sequenced. See  
 keywords for description of this sequence and for the sequence of  
 the other extremity of this insert.  
 Location/Qualifiers  
 1. 1116  
 /organism="Yarrowia lipolytica"  
 /strain="CLIB 89"  
 /db\_xref="taxon:4952"  
 /clone="AM0A016F08"  
 /clone\_1id="AM0AA"  
 /note="end : T7"

BASE COUNT 286 a 239 c 255 g 329 t 7 others

ORIGIN

Alignment Scores: 218 Length: 1116  
 Pred. No.: 69.00 Matches: 13  
 Score: 63.89% Conservative: 10  
 Percent Similarity: 63.89%

Best Local Similarity: 36.11% Mismatches: 13  
 Query Match: 23.63% Indels: 0  
 DB: 17 Gaps: 0

us-09-877-606-3 (1-52) x CNS06ROC (1-1116)

Qy 13 AsnAsnTyTrHrSerLeuIleHisSerLeuIleGluGluSerGlnAsngIngluLys 32  
 ::::: ||||| ::::: ||| ::::: ||||| ::::: |||||

Db 315 CATCACATCACCTCCACATCTCTTCCAAACGACGACGACCAAGAAAAAG 256

Qy 33 AsngIngluIngluLeuLeuAsplyStrpAlaSerLeuTPrsn 48  
 ||| ::::: |||||

Db 225 AACACGACGACAAAAAGATTCTTCAAAATGACATGCTGTACAC 208

RESULT 6  
 CNS06R23 1133 bp DNA linear GSS 05-JUL-2001  
 LOCUS T7 end of clone AM0A010G12 of library AM0A from strain CLIB 89 of  
 DEFINITION Yarrowia lipolytica, genomic survey sequence.  
 AL411409  
 AL411409.1 GI:12180799  
 GSS.  
 Yarrowia lipolytica.  
 Yarrowia lipolytica  
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 Saccharomycetales; Dipodascaceae; Yarrowia.  
 Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,  
 1 (bases 1 to 1133)  
 Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S.,  
 de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B.,  
 Malpertuy,A., Neuvéglise,C., Ozler-Kalogeropoulos,O., Potier,S.,  
 Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,  
 Wincker,P. and Weissenbach,J.  
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of  
 yeast species for molecular evolution studies  
 FEBS Lett. 487 (1), 3-12 (2000)

JOURNAL  
 MEDLINE 20584711  
 PUBMED 11152876

REFERENCE  
 AUTHORS 2 (bases 1 to 1133)  
 Casaregola,S., Neuvéglise,C., Lepingle,A., Bon,E., Feynerol,C.,  
 Artiguenave,F., Wincker,P. and Gallardin,C.  
 Genomic exploration of the hemiascomycetous yeasts: 17. Yarrowia  
 lipolytica  
 FEBS Lett. 487 (1), 95-100 (2000)

JOURNAL  
 MEDLINE 20584727  
 PUBMED 11152892

REFERENCE  
 AUTHORS 3 (bases 1 to 1133)  
 Genoscope.  
 Direct Submission  
 Submitted (07-SEP-2000) Genoscope - Centre National de Séquençage,  
 2 rue Gaston Crémieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :  
 segref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT  
 JOURNAL  
 TITLE This GSS is part of a random genomic sequencing program of thirteen  
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces  
 exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii,  
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces  
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia  
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,  
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to  
 5 kb were prepared and both extremities were sequenced. See  
 keywords for description of this sequence and for the sequence of  
 the other extremity of this insert.  
 Location/Qualifiers  
 1. 1133  
 /organism="Yarrowia lipolytica"  
 /strain="CLIB 89"  
 /db\_xref="taxon:4952"  
 /clone="AM0A010G12"  
 /clone\_1id="AM0AA"  
 /note="end : T7"  
 /note="similar to: >352)  
 /note="similar to: G21997 ( similarity to S. cerevisiae  
 Derip involved in degradation of misfolded soluble

misc\_feature

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="c70807"
/dev_stage="Adult"
/notes="Amin: colon; Vector: puc18; Site_1: Sma1; Site_2: Sma1; A mini-library was made by cloning application No. 196 from OREB75 PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

```

```

/organism="Danio rerio"
/db_xref="taxon:7955"
/clone_lib="Gong zebrafish ovary"
/sex="female"
/dev_stage="4-5 month"
/lab_host="DH10B (phage-resistant)"
/note="Organ: ovary (pooled); Vector: plasmidscript SK-;
Site:1: XhoI; Site:2: EcoRI; Poly A+ RNA was isolated from
the ovaries of 2 female adult zebrafish (4-5 month old)."

```



```

/dev.stage="4-5 month-old female adult fish"
/note="Organ: Ovary; Vector: Lambda Uni-ZAP XR; Site: 1;
EcORI: Site 2; XhoI: Poly A+ RNA was isolated from the
ovaries of 2 female adult zebrafish of 4-5 month-old and
the cDNAs were made using oligo dT primers and inserted
into Lambda ZAP vector at EcoRI (5'-cDNA) and XhoI
(3'-cDNA) sites using Stratagene's Lambda Uni-ZAP XR
cloning system."
BASE COUNT      118 a      89 c      83 g      113 t      1 others
ORIGIN

Alignment Scores:
Pred. No.:      77.9      Length:      404
Score:          67.50     Matches:      17
Percent Similarity: 48.21% Conservative: 10
Best Local Similarity: 30.36% Mismatches: 18
Query Match:    23.12%   Indels:      11
DB:             13      Gaps:      2

US-09-877-606-3 (1-52) x BM140718 (1-404)

QY      2  Histhrhrtrpreuglutrpsargluilleasnsnlyr----- 15
      |||:::      |||      |||:::
DB      34  CACAGCCTCTTCACAGATGATCAGTCAATTCGTGAGCCCTAGCAGCAGTTTCG 93
      |||:::      |||      |||:::
QY      16  ThsSerleulleHisSerleullegluIserGlnasnnglngluIlysasnglu--- 34
      ::::      ::::      ::::
DB      94  TCAAGACCTCTACAGCTCGTCAAAAGATGCACAAACAGACAGAAAGAAATTCAGA 153
      ::::      ::::      ::::
QY      35  -----GlnGluLeulleuaspIystrpIasrleu 46
      |||:::      |||:::
DB      154 AGATCGCAATGCACAGCATGTGATTGCATCATCGCTCATTCG 201

RESULT 11
LOCUS      BJ180020      508 bp      mRNA      linear      EST 24-JAN-2002
DEFINITION      BJ180020 normalized full length cDNA library, chloronemata,
caulonemata and malformed buds. Physcomitrella patens subsp. patens
cDNA clone pphb27e18 5', mRNA sequence.
ACCESSION      BJ180020
VERSION      BJ180020.1 GI:18347973
KEYWORDS      EST.
SOURCE      Physcomitrella patens subsp. patens.
ORGANISM      Physcomitrella patens subsp. patens
REFERENCE      Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
AUTHORS      Bryopsida; Funariidae; Funariales; Funariaceae; Physcomitrella.
1 (bases 1 to 508)
Fujiita,T., Shio-i,T., Seki,M., Kamiya,A., Uchiyama,I., Nishiyama,T.,
Carninci,P., Hayashizaki,Y., Shinozaki,K., Kohara,Y. and Hasebe
,M.
TITLE      Comparison of the moss Physcomitrella patens genome with flowering
JOURNAL      plants genome
COMMENT      Unpublished (2002)
Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp
A backbone of the vector is pBluescript II, that was in vivo
excised from a modified lps phage vector (Mo bi Tec, Germany). XhoI
digested-5' end of cDNA is ligated to SalI site of the vector, and
the BamHI digested-3' end including poly-A tail is ligated to BamHI
site of the vector. cDNA insert could be amplified with
conventional T7 and T3 primers. This normalized full-length cDNA
library was generated basically according to the method described
in Genome Research 10, 1617-1630 (2000), Carninci, P. et al.
Procnemata were blended by the POLYTRON, and then cultivated on
the BCD medium containing 0.5mM BA (benzylaminopurine) for 8 to 13
days under the continuous light.
FEATURES
source      1..508
Location/Qualifiers

```

```

/organism="Physcomitrella patens subsp. patens"
/db_xref="taxon:145481"
/clone_lib="pphb27e18"
/clone_lib="normalized full length cDNA library,
chloronemata, caulonemata and malformed buds"
/tissue_type="mixture of chloronemata, caulonemata and
malformed buds"
BASE COUNT      102 a      119 c      151 g      136 t
ORIGIN

Alignment Scores:
Pred. No.:      108      Length:      508
Score:          67.50     Matches:      16
Percent Similarity: 44.64% Conservative: 9
Best Local Similarity: 28.57% Mismatches: 22
Query Match:    23.12%   Indels:      9
DB:             13      Gaps:      2

US-09-877-606-3 (1-52) x BJ180020 (1-508)

QY      5  Tripleu-----GluTrpAspArgGluIleAsnAsnlyrThr 16
      |||||      |||||
DB      337  TGCGTATGTGCATGTTAGCGTCAGACAGCGGAATGACGAAGAGCTTAACAAC---CGT 393
      |||||      |||||
QY      17  SerleulleHisSerleullegluIserGlnasnnglngluIlysasngluIglu 36
      |||      ::::      |||
DB      394  TCAGGGGCTGAAGAAGAGTACACTACACAGATTCTCAAGATTGGAAGAGAGCTT 453
      |||      ::::      |||
QY      37  LeulleuGluLeuaspIystrpIasrleuTrpAsntrpIheasnlle 52
      |||||      ::::      |||
DB      454  CTGCTGCACGGGAGTGTGTTTCAGGATCCAGAACTGGTCAGATTGA 501
      |||||      ::::      |||

RESULT 12
LOCUS      BJ063369      576 bp      mRNA      linear      EST 07-DEC-2001
DEFINITION      BJ063369 NIBB Mochli normalized Xenopus tailbud library Xenopus
laevis cDNA clone XL076b18 5', mRNA sequence.
ACCESSION      BJ063369
VERSION      BJ063369.1 GI:17424660
KEYWORDS      EST.
SOURCE      African clawed frog.
ORGANISM      Xenopus laevis
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Piploidae;
Xenopodinae; Xenopus.
1 (bases 1 to 576)
Kitayama,A., Terasaka,C., Mochli,M., Ueno,N., Shin-i,T. and Kohara
,Y.
TITLE      Expressed genes in X. laevis embryo
JOURNAL      Unpublished (2001)
COMMENT      Contact: Tadasu Shin-i
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
FEATURES
source      1..576
Location/Qualifiers
/organism="Xenopus laevis"
/db_xref="taxon:8355"
/clone_lib="XL076b18"
/clone_lib="NIBB Mochli normalized Xenopus tailbud
library"
/tissue_type="whole embryo"
/dev.stage="stage 25"
/note="Vector: pBSRN3; Site: 1: NotI; Site: 2: EcoRI; cDNAs
were oligo-dT primed and directionally cloned. Staging
according to Newkooop and Faber. Library is substracted
and was constructed by N. Garrett and A.M. Zorn,
(Wellcome/CRC Institute)."
BASE COUNT      190 a      90 c      149 g      143 t      4 others
ORIGIN

```



```

/strain="SJD"
/db_xref="taxon:7955"
/clone="5627665"
/clone_lib="zebrafish STD 5 day embryo"
/tissue_type="whole embryo, 5 day"
/lab_host="DH10B"
/notes="Vector: pAMP1; Site:1: EcoRI; Site:2: NotI; First
strand cDNA synthesis was primed using oligo-dt on
magnetic beads with an additional primer
5'-ggcgccgaatacgaactacataagg-3'. Second strand
synthesis was a 3-cycle PCR using the primers
5'-ggcgccgaatacgaactacataagg-3' and
5'-aacgacgtgtaacacgacgaactc-tttttttttt-3'. cDNA
was subsequently amplified in a 7-cycle PCR with the
following primers: 5'-ggcgccgaatacgaactacataagg-3' and
5'-aacgacgtgtaacacgacgaactc-tttttttttt-3' and
a third PCR (5 cycles) and the primers
5'-caucaucauagggcgcaatacgaactacataagg-3' and
5'-cuacuacuaaagcagtgtaacacgacgaactc-3'. Ends were
treated with uracil DNA glycosylase and product with 3'
overhangs was annealed to complementary ends of pAMP1.
Insert can be excised using EcoRI and NotI. Library
constructed by Joe Barnes and Steve Johnson (Washington
University)."
BASE COUNT      130 a      99 c      99 g      126 t
ORIGIN

Alignment Scores:
Pred. No.:      107      Length:      454
Score:          67.00      Matches:      23
Percent Similarity: 42.50%      Conservative: 11
Best Local Similarity: 28.75%      Mistatches: 10
Query Match:    22.95%      Indels:      36
DB:             13      Gaps:          5

US-09-877-606-3 (1-52) x BM574215 (1-454)

QY      3  ThThThTrP-----LcGlUTrPaSPaRgLUlLeAsnAsnTYrThrSer----- 17
      Db  18  TCTACGTGGCGTACACACAGTGCAGTCTGCTTACAGCTGACACCGCTTACCT 77
QY      18  Leulle-----HisSerLeulle----- 23
      Db  78  TTGATTGGCCACACACAGCTCATTTCAAGATGATCAGTCAGTTCGTGGAGCCTA 137
QY      24  -----GluGluSerGlnAsnGlnGlnGlu 31
      Db  138  GCAGGCGATTGCTCAAGACTCTATCAGACTGTCAAAGATGCGACAAACACAGACAGAA 197
QY      32  LysAsnGlu-----GlnGluLeuLeuGluLeuAspLysTrpAlaSerLeu 46
      Db  198  AAGATTTCCAGAGATGCGCATGCGACACAGATTGGATTGGCCATCATGCGCTTCATTG 257

RESULT 15
AA438382/c
LOCUS
DEFINITION
          495 bp      mRNA      linear      EST 19-APR-2001
ACCESSION
          LD12649.5prtime LD Drosophila melanogaster embryo Bluescript
VERSION
          AA438382
KEYWORDS
          AA438382.1 GI:2792870
SOURCE
          fruit fly.
ORGANISM
          Drosophila melanogaster
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
          Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
          1 (bases 1 to 495)
          Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
          Lewis,S. and Rubin,G.M.
          BDGP/HMI Drosophila EST Project

```

```

JOURNAL
COMMENT
          Unpublished (2001)
          On Jan 19, 1998 this sequence version replaced gi:2150263.
          Contact: Stapleton, M.
          BDGP
          Lawrence Berkeley National Lab
          One Cyclotron Rd, Berkeley, CA 94720, USA
          Fax: 510 486 6798
          Email: http://www.fruitfly.org/EST, est@fruitfly.berkeley.edu
          Plate: 126 row: E column: 1
          High quality sequence stop: 481.
          Location/Qualifiers
            1. 495
              /organism="Drosophila melanogaster"
              /db_xref="BDGP EST:BDc01n11820"
              /db_xref="taxon:7227"
              /clone="LD12649"
              /clone_lib="LD Drosophila melanogaster embryo Bluescript"
              /sex="male and female"
              /dev_stage="0 to 24 hours mixed stage embryonic"
              /lab_host="SOLR"
              /note="Organ: embryo; Vector: Bluescript SK; Site:1: EcoRI
              ; Site:2: XhoI; Constructed using Stratagene ZAP-cDNA
              synthesis kit. Oligo dt-primed and directionally cloned at
              EcoRI and XhoI in Bluescript SK(+/-)"
BASE COUNT      145 a      128 c      146 g      76 t
ORIGIN

Alignment Scores:
Pred. No.:      121      Length:      495
Score:          67.00      Matches:      13
Percent Similarity: 42.22%      Conservative: 6
Best Local Similarity: 28.89%      Mistatches: 12
Query Match:    22.95%      Indels:      14
DB:             9      Gaps:          1

US-09-877-606-3 (1-52) x AA438382 (1-495)

QY      5  TrPLeuGluTrPaSPaRgLUlLeAsnAsnTYrThrSerLeulleHisSerLeulleGlu 24
      Db  143  TGGGTCCGGTGGCGGAGATTGGTCCGACACGCGACATGCCCTCATCAATGTCGTGTC--- 87
QY      25  GluSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuLeuAspLysTrpAla 44
      Db  86  -----ATCGCCTTGAGACCTTGCGCG 66
QY      45  SerLeuTrPaSPaRg 49
      Db  65  CCACCTTGCTCTGG 51

Search completed: March 25, 2003, 17:27:48
Job time : 1470.73 secs

```

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:27:06 ; Search time 60.7143 Seconds  
(without alignments)  
186.551 Million cell updates/sec

Title: US-09-877-606-4

Perfect score: 452  
Sequence: 1 NHTTWLEMDREINNTSLIH.....SKIYHIENIARIKKLIGER 85

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

A.Geneseq\_101002:\*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*
- 23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	64.2	268	19	AAV22820
2	290	64.2	268	23	ABG68291
3	290	64.2	344	23	ABB83400
4	290	64.2	420	15	AAAS3785
5	290	64.2	519	17	AAW00181
6	290	64.2	521	16	AAAR6997
7	290	64.2	853	19	AAWA3066
8	290	64.2	856	7	AAAP6031
9	290	64.2	856	14	AAAR41025
10	290	64.2	856	14	AAAR41026

11	290	64.2	856	14	AAAR41027
12	290	64.2	856	14	AAAR41028
13	290	64.2	856	14	AAAR41029
14	290	64.2	856	14	AAAR41030
15	290	64.2	856	14	AAAR41031
16	290	64.2	856	14	AAAR41032
17	290	64.2	856	21	AAAY7072
18	290	64.2	856	22	AAAB45697
19	278	61.5	177	23	ABBR83401
20	278	61.5	177	23	ABBR83413
21	277	61.3	56	21	ABBI4532
22	277	61.3	56	22	AAU70191
23	277	61.3	56	22	AAAG63860
24	277	61.3	138	21	AAAG69792
25	277	61.3	150	9	AAAP80745
26	277	61.3	192	20	AAAY24098
27	277	61.3	233	17	AAAG93199
28	277	61.3	268	19	AAAY22821
29	277	61.3	268	23	ABG68292
30	277	61.3	268	23	ABG68293
31	277	61.3	268	23	AAAR25787
32	277	61.3	342	13	ABBR83410
33	277	61.3	344	23	ABBR83410
34	277	61.3	345	21	AAAB14536
35	277	61.3	345	22	AAAG63863
36	277	61.3	393	15	AAAR52687
37	277	61.3	393	15	AAAR45445
38	277	61.3	395	10	AAAP91956
39	277	61.3	709	10	AAAP95781
40	277	61.3	751	10	AAAP90604
41	277	61.3	851	9	AAAP80967
42	277	61.3	854	21	AAAB10697
43	277	61.3	854	22	AAAB10053
44	277	61.3	854	22	AAAB61199
45	277	61.3	854	22	AAAB67277

#### ALIGNMENTS

RESULT 1	
AAV22820	AAV22820 standard; Protein; 268 AA.
XX	XX
AC	AAV22820;
XX	XX
DT	19-AUG-1999 (first entry)
XX	XX
DE	SEQ ID NO. 16 from WO9820036.
XX	XX
KW	HIV, gp41 protein; constrained helical peptide; HIV infection;
KW	vaccine; antibody; viral membrane fusion; viral infectivity;
KW	ligand affinity purification; protein A replacement;
KW	immunoglobulin purification; epitope mimic.
XX	XX
OS	Human immunodeficiency virus.
XX	XX
FN	WO9820036-A1.
XX	XX
PD	14-MAY-1998.
XX	XX
PE	05-NOV-1997; 97WO-US20069.
XX	XX
PR	16-JUN-1997; 97US-0876698.
XX	XX
PR	06-NOV-1996; 96US-0743698.
XX	XX
PA	(GETH ) GENENTECH INC.
XX	XX
PI	Bristled A, Judice JK, McDowell RS, Starovassnik MA;
XX	XX
PI	Wells JA;
XX	XX
DR	WPI; 1998-286866/25.
XX	XX

Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Selectively deglyc  
Wild type HIV-1 HX  
HIV gp41 LAI prote  
HIV gp41 LAI prote  
HIV-1 isolate LAI  
HIV viral envelope  
Amino acid sequenc  
HIV-1 gp41 envelop  
Sequence of AAs 60  
Recombinant gp41 p  
SEQ ID NO. 17 from  
SEQ ID NO. 18 from  
Envelope protein g  
ENV9. Synthetic.  
HIV gp41 LAI prote  
HIV-1 isolate LAI  
Amino acid sequenc  
gp41 HIV envelope  
Sequence encoded b  
HIV p41 gene with  
Fusion protein of  
HIV-1 env segment  
HIV protein HT6.  
HIV-1 coat protein  
HIV gp41 DNA SEQ I  
Protein encoded by



PT vaccines against human immune deficiency virus, mimics the intermediate  
 PT state of the native polypeptide  
 XX  
 PS Disclosure; Page 29; 29pp; French.  
 XX  
 CC The present invention relates to novel mutant HIV gp41 LAI proteins  
 CC (AB83411-AB83420). The mutants can form a structure corresponding to,  
 CC or mimicking, the intermediate state of gp41. The mutants, or their  
 CC conjugate with a carrier, or vectors containing nucleic acid that encode  
 CC them, are used in vaccines for treatment or prevention of infection by  
 CC HIV. The present sequence is wild-type HIV gp41 LAI protein, which  
 CC was used to generate the mutants of the invention.  
 XX  
 SQ Sequence 344 AA;  
 Query Match 64.2%; Score 290; DB 23; Length 344;  
 Best Local Similarity 98.1%; Pred. No. 1.9e-20;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQLELDKWSLMMNFI 52  
 DB 113 NHTTWMDREINNTSLIHSLSIESQNOEKNEQLELDKWSLMMNFI 164  
 RESULT 4  
 AAR53785  
 ID AAR53785 standard; Protein; 420 AA.  
 AC AAR53785;  
 DT 28-DEC-1994 (first entry)  
 XX  
 DE Translation of HIV-1 in the region encoding the gp41 polypeptide.  
 XX  
 KM Human immunodeficiency virus; HIV-1; AIDS; lentivirus; TM protein;  
 KM extracellular protein; transmembrane protein; gp41.  
 XX  
 OS HIV-1.  
 FH Key Location/Qualifiers  
 FT Protein 43..387  
 FT /label= gp41  
 XX  
 PN WO9412533-A.  
 PD 09-JUN-1994.  
 XX  
 PF 12-JAN-1993; 93WO-US00212.  
 XX  
 PR 23-NOV-1992; 92US-0979975.  
 XX  
 PA (HARD ) HARVARD COLLEGE.  
 XX  
 PI Essex ME, Lee TH, Yu X;  
 XX  
 DR WPI; 1994-200197/24.  
 DR N-PSDB; AA066275.  
 XX  
 PT Method for treating HIV patients - comprises administration of  
 PT mutated gp41 polypeptide  
 XX  
 PS Claim 11; Fig 1; 54pp; English.  
 XX  
 CC The inventors claim a method of treating a patient infected with HIV  
 CC by administering a mutated gp41 polypeptide or a therapeutic  
 CC composition comprising nucleic acid encoding the mutant gp41  
 CC polypeptide in an expressible genetic construction. The mutant gp41  
 CC polypeptide contains a deletion of at least one AA in at least one  
 CC of the following regions of wild type gp41 (AAR53783): AAs 844-856;  
 CC 814-856; 796-856; 776-856; 753-856; or 710-856, effective to either  
 CC disrupt viral replication or HIV or disrupt the assembly of viral  
 CC Env proteins in an HIV infected cell. AA066275 corresp. to bps 7631-  
 CC 8890 of wt HIV-1. X in the AA sequence represents the posn. of a

CC stop codon in AA066275.  
 XX  
 SQ Sequence 420 AA;  
 Query Match 64.2%; Score 290; DB 15; Length 420;  
 Best Local Similarity 98.1%; Pred. No. 2.4e-20;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQLELDKWSLMMNFI 52  
 DB 155 NHTTWMDREINNTSLIHSLSIESQNOEKNEQLELDKWSLMMNFI 206  
 RESULT 5  
 AAW00181  
 ID AAW00181 standard; Protein; 519 AA.  
 AC AAW00181;  
 DT 12-NOV-1996 (first entry)  
 XX  
 DE HIV-1 env protein.  
 XX  
 KM Poliovirus; encapsidation; capsid; vaccine; genetic immunisation;  
 KM HIV-1; human immunodeficiency virus type 1; env protein.  
 XX  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO9625173-A1.  
 PD 22-AUG-1996.  
 XX  
 PF 13-FEB-1996; 96WO-US01895.  
 XX  
 PR 15-FEB-1995; 95US-0389459.  
 XX  
 PA (UABR-) UAB RES FOUND.  
 XX  
 PI Ansardl DC, Morrow CD, Porter DC;  
 XX  
 DR WPI; 1996-393136/39.  
 DR N-PSDB; AAT33295.  
 XX  
 PT Encapsidation of recombinant polio:virus nucleic acid for use in  
 PT vaccines - using a polio:virus nucleic acid which lacks the P1  
 PT capsid region and an expression system which provides the region  
 XX  
 PS Disclosure; Page 61-63; 108pp; English.  
 XX  
 CC CDNA sequences (AAT33293-95) respectively code for the gag, pol and  
 CC env proteins (AAW00179-81) of HIV-1. They can be used to substitute  
 CC the P1 capsid gene of poliovirus in recombinant poliovirus nucleic  
 CC acids (rPNAs). Such rPNAs are encapsidated by introduction into  
 CC a host cell together with a vaccinia virus or plasmid vector  
 CC encoding the poliovirus P1 capsid precursor protein. Encapsidated  
 CC rPNAs are useful for genetic immunisation, stimulating an immune  
 CC response to the HIV-1 protein.  
 XX  
 SQ Sequence 519 AA;  
 Query Match 64.2%; Score 290; DB 17; Length 519;  
 Best Local Similarity 98.1%; Pred. No. 3.1e-20;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQLELDKWSLMMNFI 52  
 DB 420 NHTTWMDREINNTSLIHSLSIESQNOEKNEQLELDKWSLMMNFI 471  
 RESULT 6  
 AAR69997  
 ID AAR69997 standard; Protein; 521 AA.

```

AC AAR69997;
XX
XX 13-SEP-1995 (first entry)
XX
XX HIV-1 env protein.
XX
XX Encapsulation; Human immunodeficiency virus; type 1; immunogenic;
XX poliovirus.
XX
XX Human immunodeficiency virus type 1.
XX
XX CA2125344-A.
XX
XX 02-JAN-1995.
XX
XX 07-JUN-1994; 94CA-2125344.
XX
XX 01-JUL-1993; 93US-0087009.
XX
XX (UABR-) UAB RES FOUND.
XX
XX Morrow CD;
XX
XX WPI; 1995-099021/14.
XX
XX N-PSDB; AAR60575.
XX
XX Method for encapsulating recombinant polio:virus nucleic acid
XX useful for providing compsns. to stimulate immune response.
XX
XX Disclosure; Page 38; 62pp; English.
XX
XX The sequence is that of the HIV-1 env protein. The DNA encoding
XX such protein is used in a method (claimed) to encapsulate
XX poliovirus cDNA to make it more immunogenic.
XX See also AAR6995-6.
XX
XX Sequence 521 AA;
XX
XX Query Match 64.2%; Score 290; DB 16; Length 521;
XX Best Local Similarity 98.1%; Pred. No. 3.1e-20;
XX Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 NHTTWLEMDREINNYTSLHSLIEESONQOEKNEOLLELDKASLWNWFI 52
DB 420 NHTTWLEMDREINNYTSLHSLIEESONQOEKNEOLLELDKASLWNWFI 471

```

RESULT 7

```

AAM43066
ID AAM43066 standard; peptide; 853 AA.
XX
XX AAM43066;
XX
XX 11-SEP-1998 (first entry)
XX
XX HIV-1 gp120 protein fragment from isolate HXB2.
XX
XX gp120 protein; purification; fractionation; ion exchange; chromatography;
XX binding affinity; CD4; hydrophobic interaction; size exclusion; vaccine.
XX
XX Human immunodeficiency virus type 1.
XX
XX US5696238-A.
XX
XX 09-DEC-1997.
XX
XX 11-MAY-1995; 95US-0439286.
XX
XX 20-AUG-1991; 91US-0684963.
XX
XX 16-AUG-1993; 93US-0109002.
XX
XX 09-MAY-1994; 94US-0240073.
XX
XX 11-MAY-1995; 95US-0439286.
XX

```

```

PA (CHIR ) CHIRON CORP.
XX
XX Halgwood NL, Scandella C;
XX
XX WPI; 1998-041353/04.
XX
XX Purification of HIV gp120 - using chromatographic methods
XX
XX Disclosure; Fig 2A-W; 53pp; English.
XX
XX AAM43066-W43080 are fragments of the gp120 protein from different human
XX immunodeficiency virus type I (HIV-1) isolates. These proteins are used
XX in a novel method for purifying HIV gp120 so as to provide a purified
XX gp120 glycoprotein having protein/protein binding properties
XX substantially identical to natural viral HIV gp120. The method involves
XX fractionating a crude gp120 preparation containing full-length,
XX glycosylated gp120 using ion exchange chromatography so as to provide a
XX first collection of fractions. A fraction from the first collection is
XX selected that exhibits specific binding affinity for CD4 peptide,
XX thereby producing a first fractionated material. The first fractionated
XX material is fractionated by hydrophobic interaction chromatography so as
XX to provide a second collection of fractions from which a second
XX collection is selected that exhibits specific binding affinity for CD4
XX peptide. This second fraction is fractionated by size exclusion
XX chromatography so as to provide a third collection of fractions
XX exhibiting specific binding affinity for CD4 peptide, thereby providing
XX the purified gp120. The purified gp120 can be used for antibody
XX production and in vaccines.
XX
XX Sequence 853 AA;
XX
XX Query Match 64.2%; Score 290; DB 19; Length 853;
XX Best Local Similarity 98.1%; Pred. No. 5.4e-20;
XX Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 NHTTWLEMDREINNYTSLHSLIEESONQOEKNEOLLELDKASLWNWFI 52
DB 621 NHTTWLEMDREINNYTSLHSLIEESONQOEKNEOLLELDKASLWNWFI 672

```

RESULT 8

```

ID AAP60131
XX
XX AAP60131 standard; Protein; 856 AA.
XX
XX AAP60131;
XX
XX 26-JUN-1991 (first entry)
XX
XX Sequence of the AIDS envelope protein.
XX
XX AIDS; HIV; LAV; HTLV-III; vaccine; antibody; epitope; antigen;
XX diagnosis.
XX
XX HTLV-III.
XX
XX EPI99301-A.
XX
XX 29-OCT-1986.
XX
XX 18-APR-1986; 86EP-0105371.
XX
XX 19-APR-1985; 85US-0725021.
XX
XX (HOFF ) HOFFMANN-LA ROCHE AG.
XX (USGO ) US GOVERNMENT.
XX (HEAL-) DEPT. HEALTH & HUMAN SERV.
XX (USDH ) US DEPT HEALTH & HUMAN.
XX
XX Crowl RM, Gallo RC, Reddy EP, Shaw GM, Wong-Staal FY;
XX WPI; 1986-286067/44.
XX
XX N-PSDB; AAN60128.
XX

```

PT Envelope protein of acquired immune deficiency syndrome virus -  
PT useful for improved testing of human blood for antibodies against  
PT virus and as antigen for vaccines  
XX  
PS Claim 33: Fig 6A; 46pp; English.  
XX  
CC An expression vector contg. a gene (AAN60128) coding for an envelope  
CC protein of an AIDS virus, and the envelope protein of an AIDS virus  
CC (AAP60131) are claimed. The vector is pref. a member of the pEV/env  
CC family, e.g. pEV1.2 or 3/env 44-640 or 205-640.  
XX  
SQ Sequence 856 AA:  
  
Query Match 64.2%; Score 290; DB 7; Length 856;  
Best Local Similarity 98.1%; Pred. No. 5.5e-20;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NHTTWLEMDREINNYSLIHSLEESQNOEKNEDELLELDKNASLMMWPMNI 52  
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:  
DB 624 NHTTWLEMDREINNYSLIHSLEESQNOEKNEDELLELDKNASLMMWPMNI 675  
  
RESULT 9  
AAR41025  
ID AAR41025 standard; protein; 856 AA.  
XX  
AC AAR41025;  
XX  
DT 23-MAR-1994 (first entry)  
XX  
DE Selectively deglycosylated HIV-1 HXB2 env gp160 mutein C4.  
XX  
XX N-linked glycosylation; gp120; human immunodeficiency virus; type 1;  
KM envelope glycoprotein; vaccine; site-directed mutagenesis.  
XX  
OS Human immunodeficiency virus type 1 (HXB2 strain).  
XX  
FH Key Location/Qualifiers  
FH Peptide 1..30  
FT /label= signal\_peptide  
FT 31..511  
FT /label= gp120  
FT /note= "exterior membrane glycoprotein"  
FT 512..856  
FT /label= gp41  
FT /note= "transmembrane glycoprotein"  
FT 88  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 136  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 141  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 156  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 160  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 186  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 197  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 230  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 234  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
XX

FT Modified-site 241  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 262  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 276  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 289  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 295  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 301  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 332  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 339  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 356  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 392  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 448  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 611  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 616  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 624  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 637  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 674  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 750  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 816  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 836  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 866  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 897  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 906  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 943  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 963  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 963  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
XX

PD 16-SEP-1993.  
 XX 24-FEB-1993; 93WO-US01598.  
 XX 13-MAR-1992; 92US-0850770.  
 XX (HARD ) HARVARD COLLEGE.  
 XX Essex ME, Lee C, Lee T, Lee W;  
 XX WPI; 1993-303140/38.  
 DR  
 PT Compn. contg. selectively de-glycosylated HIV-1 envelope protein  
 PT - shows improved protective immune response  
 PS Claim 7, Page 15 and 23; 45pp; English.  
 XX  
 CC Mutant gp160 env protein C4 is specifically claimed. The  
 CC amino acid sequence of the mutain does not appear in the  
 CC specification; sequence R41025 has been derived from the HIV-1 HXB2  
 CC isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIV1X,  
 CC Acc.#: P04578) and the description of C4 mutain given in the  
 CC specification. The combination of changes made to N-linked  
 CC glycosylation sites in the C-terminal region of gp120 does not prevent  
 CC viral infectivity but the resultant selective deglycosylation enables  
 CC an immune response to be elicited by the mutain.  
 XX  
 SQ Sequence 856 AA;  
 Query Match 64.2%; Score 290; DB 14; Length 856;  
 Best Local Similarity 98.1%; Pred. No. 5.5e-20;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NHTTLMEDREINNYTSLHSLIESQNOQEKNEDELTDKMSLWMWENI 52  
 ID 624 NHTTLMEDREINNYTSLHSLIESQNOQEKNEDELTDKMSLWMWENI 675  
 RESULT 10  
 AAR41026  
 ID AAR41026 standard; protein; 856 AA.  
 XX  
 AC AAR41026;  
 XX  
 DT 23-MAR-1994 (first entry)  
 XX  
 DE Selectively deglycosylated HIV-1 HXB2 env gp160 mutain C5.  
 XX  
 KM N-linked glycosylation; gp120; human immunodeficiency virus; type 1;  
 KM envelope glycoprotein; vaccine; site-directed mutagenesis.  
 XX  
 OS Human immunodeficiency virus type 1 (HXB2 strain).  
 XX  
 XX Key Location/Qualifiers  
 FH 1.30  
 FH Peptide /label= signal-peptide  
 FT 31..511  
 FT Protein /label= gp120  
 FT /note= "exterior membrane glycoprotein"  
 FT 512..856  
 FT Protein /label= gp41  
 FT /note= "transmembrane glycoprotein"  
 FT 88  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 136  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 141  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 156  
 FT Modified-site /label= N-linked\_glycosylation\_site

FT Modified-site /note= "putative"  
 FT /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 186  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 197  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 230  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 234  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 241  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 262  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 276  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 289  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 295  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 301  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 332  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 339  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 356  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 448  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 611  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 616  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 624  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 637  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 674  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 750  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 816  
 FT Modified-site /label= N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT 386  
 FT Misc-difference /note= "putative"  
 FT 392  
 FT Misc-difference /note= "putative N-linked glycosylation site  
 eliminated by substitution of Asn by  
 Gln at position 386"

FT eliminated by substitution of Asn by  
FT Gln at position 392"  
FT Misc-difference 397  
FT /note= "putative N-linked glycosylation site  
FT eliminated by substitution of Asn by  
FT Gln at position 397"  
FT Misc-difference 406  
FT /note= "putative N-linked glycosylation site  
FT eliminated by substitution of Asn by  
FT His at position 406"  
FT Misc-difference 463  
FT /note= "putative N-linked glycosylation site  
FT eliminated by substitution of Asn by  
FT His at position 463"  
FT W09317705-A.  
FT 16-SEP-1993.  
FT 24-FEB-1993; 93MO-US01598.  
FT 13-MAR-1992; 92US-0850770.  
FT (HARD ) HARVARD COLLEGE.  
FT Essex ME, Lee C, Lee T, Lee W;  
FT WPI, 1993-303140/38.  
FT Compsn. contg. selectively de-glycosylated HIV-1 envelope protein  
FT - shows improved protective immune response  
FT Claim 7; Page 15 and 23; 45pp; English.  
FT  
FT PS Mutant gp160 env protein C5 is specifically claimed. The  
FT CC amino acid sequence of the mutein does not appear in the  
FT CC specification; sequence R41026 has been derived from the HIV-1 HXB2  
FT CC isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIYX,  
FT CC Acc. #: P04578) and the description of C5 mutein given in the  
FT CC specification. The combination of changes made to N-linked  
FT CC glycosylation sites in the C-terminal region of gp120 does not prevent  
FT CC viral infectivity but the resultant selective deglycosylation enables  
FT CC an immune response to be elicited by the mutein.  
FT SQ Sequence 856 AA:  
  
Query Match 64.2%; Score 290; DB 14; Length 856;  
Best Local Similarity 98.1%; Pred. No. 5,5e-20;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 NHTTWMEMREINNTSLHSLIEESONQOEKNEQELLELDKWSIMNPFNI 52  
DB 624 NHTTWMEMREINNTSLHSLIEESONQOEKNEQELLELDKWSIMNPFNI 675  
  
RESULT 11  
ID AAR41027  
XX AAR41027 standard: protein; 856 AA.  
AC AAR41027;  
DT 23-MAR-1994 (first entry)  
XX  
XX Selectively deglycosylated HIV-1 HXB2 env gp160 mutein C6.  
DE  
XX  
XX N-linked glycosylation; gp120; human immunodeficiency virus; type 1;  
KW envelope glycoprotein; vaccine; site-directed mutagenesis.  
XX  
OS Human immunodeficiency virus type 1 (HXB2 strain).  
XX  
FH Key Location/Qualifiers  
FT 1..30  
FT Peptide /label= signal\_peptide

FT Protein  
FT 31..511  
FT /label= gp120  
FT /note= "exterior membrane glycoprotein"  
FT 512..856  
FT /label= gp41  
FT /note= "transmembrane glycoprotein"  
FT 88  
FT /label= N-linked glycosylation\_site  
FT /note= "putative"  
FT 136  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 141  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 156  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 160  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 186  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 197  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 230  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 234  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 241  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 262  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 276  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 289  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 295  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 301  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 332  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 339  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 356  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 611  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 616  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 624  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 637  
FT /label= N-linked\_glycosylation\_site  
FT /note= "putative"  
FT 674  
FT Modified-site

Query Match	Best Local Similarity	Score	DB	Length
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;	64.2%;	290;	DB 14;	856;
1 NHTTLEMDREINNTSLIHSLEESQDQENDEBLLFELDKWASLWMPFI 52	98.1%;	Pred. No. 5.5e-20;		
624 NHTTLEMDREINNTSLIHSLEESQDQENDEBLLFELDKWASLWMPFI 675				

RESULT 12	
AA041028	
ID	AA041028 standard; protein; 856 AA.
XX	
AC	AA041028;
XX	
DT	23-MAR-1994 (first entry)
XX	
DE	Selectively deglycosylated HIV-1 HXB2 env gp160 mutain Q.
XX	
KW	N-linked glycosylation; gp120; human immunodeficiency virus envelope glycoprotein; vaccine; site-directed mutagenesis
XX	
OS	Human immunodeficiency virus type 1 (HXB2 strain).
XX	
PH	Key
FT	1..30
FT	/label= "signal_peptide
FT	31..511
FT	/label= gp120
FT	/note= "exterior membrane glycoprotein"
FT	512..856
FT	/label= gp41
FT	/note= "transmembrane glycoprotein"
FT	88
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	136
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	141
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	156
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	160
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	186
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	197
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	230
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	234
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	241
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	262
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	276
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	289
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	295
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	301
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	392
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	448



FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	Modified-site
FT	611
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	Modified-site
FT	616
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	Modified-site
FT	624
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	Modified-site
FT	637
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	Modified-site
FT	674
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	Modified-site
FT	750
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	Modified-site
FT	816
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	Misc-difference
FT	332
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 332"
FT	Misc-difference
FT	339
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 339"
FT	Misc-difference
FT	356
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 356"
FT	Misc-difference
FT	386
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
FT	Misc-difference
FT	397
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FT	Misc-difference
FT	406
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FT	Misc-difference
FT	463
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
XX	W09317705-A.
PD	16-SEP-1993.
XX	24-FEB-1993:
PR	93WC-US01598.
XX	13-MAR-1992;
XX	92US-0850770.
PA	(HARD ) HARVARD COLLEGE.
XX	Essex ME, Lee C, Lee T, Lee W;
PL	WP1; 1993-303140/38.
DR	Compsn. contg. selectively de-glycosylated HIV-1 envelope protein
XX	- shows improved protective immune response
PT	Claim 7; Page 18, 23 and 26; 45pp: English.
CC	Mutant gp160 env protein Q is specifically claimed. The amino acid sequence of the mutain does not appear in the

CC	specification: sequence RA1028 has been derived from the HIV-1 HXB2
CC	Isolate gp160 env sequence on the SWISS-Prot database (ID = ENVSH1YIX,
CC	Acc.# : P04578) and the description of Q mutin given in the
CC	specification. The combination of changes made to N-linked
CC	glycosylation sites in the C-terminal region of gp120 does not prevent
CC	viral infectivity but the resultant selective deglycosylation enables
CC	an immune response to be elicited by the mutin.
CC	
XX	Sequence 856 AA:
SQL	
Query Match	64.2%; Score 290; DB 14; Length 856;
Best Local Similarity	98.1%; Pred. No. 5..5e-20;
Matches 51; Conservative	1; Mismatches 0; Indels 0; Gaps 0;
QY	1 NHTTWLEMDREINNTSLIHSLEESQNOQEKNEDELLELDKWKASIMNPFNI 52
DB	624 NHTTWLEMDREINNTSLIHSLEESQNOQEKNEDELLELDKWKASIMNPFNI 675
RESULT 13	
AAK41029	
ID	AAK41029 standard; protein; 856 AA.
XX	
AC	AAK41029;
XX	
DT	23-MAR-1994 (first entry)
XX	
DE	Selectively deglycosylated HIV-1 HXB2 env gp160 mutin R.
XX	
KW	N-linked glycosylation; gp120; human immunodeficiency virus; type 1;
KM	envelope glycoprotein; vaccine; site-directed mutagenesis.
XX	
OS	Human immunodeficiency virus type 1 (HXB2 strain).
XX	
FH	Key Location/Qualifiers
FT	Peptide 1..30
FT	/label= "signal_peptide
FT	31..511
FT	/label= "gp120
FT	/note= "exterior membrane glycoprotein"
FT	512..856
FT	/label= "gp41
FT	/note= "transmembrane glycoprotein"
FT	88
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	136
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	141
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	156
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	160
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	186
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	197
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	230
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	234
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"
FT	241
FT	/label= "N-linked_glycosylation_site
FT	/note= "putative"



FT	Modified-site	160	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	186	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	197	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	230	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	234	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	241	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	262	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	276	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	289	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	295	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	356	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	392	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	448	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	611	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	616	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	624	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	637	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	674	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	730	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	816	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Modified-site	301	/label= "N-linked_glycosylation_site
FT			/note= "putative"
FT	Misc-difference	332	"putative N-linked glycosylation site
FT			eliminated by substitution of Asn by
FT			His at position 301"
FT	Misc-difference	339	"putative N-linked glycosylation site
FT			eliminated by substitution of Asn by
FT			His at position 332"
FT	Misc-difference	/note=	"putative N-linked glycosylation site
FT			eliminated by substitution of Asn by
FT			Gln at position 339"
FT	Misc-difference	386	

FT	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
FT	Misc-difference 397	/note=
FT		"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FT	Misc-difference 406	/note=
FT		"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FT	Misc-difference 463	/note=
FT		"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
PN	WO9317705-A.	
XX	16-SEP-1993.	
XX	PD	
XX	24-FEB-1993;	93WO-US01598.
PF		
XX	13-MAR-1992;	92US-0850770.
PR		
XX	(HARD ) HARVARD COLLEGE.	
PA		
XX	Essex ME, Lee C, Lee T, Lee W;	
XX	WPI: 1993-303140/38.	
DR		
XX	Compsn. conty. selectively de-glycosylated HIV-1 envelope protein	
PT	- shows improved protective immune response	
XX	Claim 7; Page 18, 23 and 26; 45pp; English.	
PS		
XX		
CC	Mutant gp160 env protein S is specifically claimed. The	
CC	amino acid sequence of the mutcin does not appear in the	
CC	specification; sequence R41030 has been derived from the HIV-1 HXB2	
CC	isolate gp160 env sequence on the SWISS-PROT database (ID = ENV\$HIV1X,	
CC	Acc.#: P04578) and the description of S mutcin given in the	
CC	specification. The combination of changes made to N-linked	
CC	glycosylation sites in the C-terminal region of gp120 does not prevent	
CC	vital infectivity but the resultant selective deglycosylation enables	
CC	an immune response to be elicited by the mutcin.	
XX		
SQ	Sequence 856 AA:	
Query Match	64.2%;	Score 290; DB 14; Length 856;
Best Local Similarity	98.1%;	Pred. No. 5.5e-20;
Matches 51; Conservative 1;	Mismatches 0;	Indels 0; Gaps 0;
OY	1	NHTTLEMDREINNTSLIHSLSIESONQEKNEDELLELDKWSLMMWFNI 52
Ddb	624	NHTTMENDREINNTSLIHSLSIESONQEKNEDELLLELDKWSLMMWFNI 675
RESULT 15		
ID	AAR41031	
XX	AAR41031 standard; protein: 856 AA.	
AC	AAR41031;	
XX		
DT	23-MAR-1994 (first entry)	
XX		
DE	Selectively deglycosylated HIV-1 HXB2 env gp160 mutcin T.	
XX		
XX	N-linked glycosylation; gp120; human immunodeficiency virus; type 1;	
KM	envelope glycoprotein; vaccine; site-directed mutagenesis.	
XX		
OS	Human immunodeficiency virus type 1 (HXB2 strain).	
XX		
Key	Location/Qualifiers	
TH	Peptide 1..30	

FT		/label= "signal_peptide
FT	Protein	31..511
FT		/label= gp120
FT		/note= "exterior membrane glycoprotein"
FT	Protein	512..856
FT		/label= gp41
FT		/note= "transmembrane glycoprotein"
FT	Modified-site	88
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	136
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	141
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	156
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	160
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	186
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	197
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	230
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	234
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	241
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	262
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	276
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	295
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	301
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	356
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	392
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	448
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	611
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	616
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	624
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	637
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"
FT	Modified-site	674
FT		/label= "N-linked_glycosylation_site
FT		/note= "putative"

FT	Modified-site	750	/label= N-linked glycosylation-site
FT		/note=	"putative"
FT	Modified-site	816	/label= N-linked glycosylation-site
FT		/note=	"putative"
FT	Misc-difference	289	/note=
FT			"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 289"
FT	Misc-difference	332	/note=
FT			"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 332"
FT	Misc-difference	339	/note=
FT			"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 339"
FT	Misc-difference	386	/note=
FT			"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
FT	Misc-difference	397	/note=
FT			"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FT	Misc-difference	406	/note=
FT			"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FT	Misc-difference	463	/note=
FT			"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
PN	W09317705-A.		
PD	16-SEP-1993.		
PE	24-FEB-1993;	93WO-US01598.	
PR	13-MAR-1992;	92US-0850770.	
PA	(HARD ) HARVARD COLLEGE.		
PI	Essex ME, Lee C, Lee T, Lee W;		
DR	WPI; 1993-303140/38.		
PT	Compsn. contg. selectively de-glycosylated HIV-1 envelope protein		
PT	- Shows Improved protective Immune response		
PS	Claim 7; Page 18, 23 and 26; 45pp; English.		
CC	Mutant gp160 env protein T is specifically claimed. The		
CC	amino acid sequence of the muteln does not appear in the		
CC	specification; sequence R41031 has been derived from the HIV-1 HXB2		
CC	isolate gp160 env sequence on the SWISS-PROT database (ID = ENV5HIV1X,		
CC	Acc.#: P04578) and the description of T muteln given in the		
CC	specification. The combination of changes made to N-linked		
CC	glycosylation sites in the C-terminal region of gp120 does not prevent		
CC	viral infectivity but the resultant selective deglycosylation enables		
CC	an immune response to be elicited by the muteln.		
SO	Sequence	856 AA;	
Query Match	Best Local Similarity	64.2%; Score 290; DB 14;	Length 856;
Matches	51; Conservative	98.1%; Pred. No. 5.5e-20;	
		1; Mismatches	0; Indels
			0; Gaps
OY	1 NHTTWLEMDREINNTSLIHSIIIESQNDKEQDELLELDKWSIIMWNT	52	
DB	624 NHTTWLEMDREINNTSLIHSIIIESQNDKEQDELLELDKWSIIMWNT	675	

Search completed: March 18, 2003, 14:33:25  
Job time : 61.7143 secs

---

**THIS PAGE BLANK (REPT)**

GenCore version 5.1.4\_p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:33:36 ; Search time 20.5172 Seconds  
(without alignments)  
121.895 Million cell updates/sec

Title: US-09-877-606-4

Perfect score: 452  
Sequence: 1 NHTTWLEMDREINNYTSLIH.....SKIVHIENIARIKRLIGER 85

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCRMUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	64.2	138	4	US-09-570-921-20
2	290	64.2	138	4	US-09-570-921-21
3	290	64.2	268	4	US-08-965-056-16
4	290	64.2	345	4	US-09-272-342B-7
5	290	64.2	519	1	US-08-589-446-8
6	290	64.2	519	1	US-08-444-882-8
7	280	64.2	519	2	US-08-459A-8
8	290	64.2	519	3	US-08-987-867A-8
9	290	64.2	856	2	US-07-916-098A-2
10	290	64.2	856	4	US-09-337-387-11
11	278	61.3	839	4	US-08-472-240A-10
12	277	61.3	138	4	US-09-570-921-22
13	277	61.3	138	4	US-09-570-921-24
14	277	61.3	138	4	US-09-570-921-26
15	277	61.3	138	4	US-09-570-921-58
16	277	61.3	237	3	US-08-388-353-641
17	277	61.3	237	3	US-08-488-551B-641
18	277	61.3	268	4	US-08-965-056-17
19	277	61.3	268	4	US-08-965-056-18
20	277	61.3	282	5	PCT-US95-1335-1
21	277	61.3	338	3	US-08-486-099-90
22	277	61.3	338	3	US-08-360-107A-100
23	277	61.3	338	3	US-08-484-223B-90
24	277	61.3	338	3	US-08-919-597-90
25	277	61.3	338	3	US-08-475-668A-90
26	277	61.3	338	3	US-08-485-551A-90
27	277	61.3	338	3	US-08-471-913A-90

28	277	61.3	338	4	US-08-485-264A-90	Sequence 90, Appl
29	277	61.3	338	4	US-08-474-349A-90	Sequence 90, Appl
30	277	61.3	338	4	US-08-255-208A-26	Sequence 26, Appl
31	277	61.3	345	4	US-08-817-441-49	Sequence 49, Appl
32	277	61.3	615	4	US-09-257-490-11	Sequence 11, Appl
33	277	61.3	826	1	US-08-375-510-2	Sequence 2, Appl
34	277	61.3	826	2	US-08-487-657-2	Sequence 2, Appl
35	277	61.3	854	4	US-09-309-572-23	Sequence 23, Appl
36	277	61.3	856	4	US-09-124-900-9	Sequence 9, Appl
37	277	61.3	861	1	US-08-127-499A-14	Sequence 14, Appl
38	277	61.3	861	1	US-08-482-847-14	Sequence 14, Appl
39	277	61.3	861	4	US-07-956-483-10	Sequence 10, Appl
40	277	61.3	861	4	US-07-956-483-16	Sequence 16, Appl
41	277	61.3	861	4	US-08-472-240A-1	Sequence 1, Appl
42	277	61.3	861	4	US-08-472-240A-7	Sequence 7, Appl
43	277	61.3	861	4	US-08-817-441-103	Sequence 103, App
44	277	61.3	863	3	US-08-463-210-11	Sequence 11, Appl
45	277	61.3	880	2	US-08-788-815-7	Sequence 7, Appl

## ALIGNMENTS

```

RESULT 1
US-09-570-921-20
; Sequence 20, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; FILE REFERENCE: 106213
; CURRENT FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US/09/570,921
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Human
US-09-570-921-20
; Sequence 21, Application US/09570921
; Patent No. 6455265
; GENERAL INFORMATION:
; APPLICANT: SERRES, PIERRE-FRANCOIS
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
; FILE REFERENCE: 106213
; CURRENT FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: US/09/570,921
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: PCT/FR98/02447
; PRIOR FILING DATE: 1997-11-17
; NUMBER OF SEQ ID NOS: 144
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
;
Query Match      64.2%; Score 290; DB 4; Length 138;
Best Local Similarity 98.1%; Pred. No. 2e-22;
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 NHTTWLEMDREINNYTSLIHSLIEESONOEKNEOELLEDKWASLMMWFI 52
DB 85 NHTTWLEMDREINNYTSLIHSLIEESONOEKNEOELLEDKWASLMMWFI 136

```

LENGTH: 138  
TYPE: PRT  
ORGANISM: Human  
US-09-570-921-21

Query Match 64.2%; Score 290; DB 4; Length 138;  
Best Local Similarity 98.1%; Pred. No. 2e-22;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKWSLWMPNI 52  
|||||  
Db 85 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKWSLWMPNI 136

RESULT 3  
US-08-965-056-16  
Sequence 16, Application US/08965056  
Patent No. 6271198  
GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted  
APPLICANT: J. Kevin Judice  
APPLICANT: Robert S. McDowell  
APPLICANT: J. Christopher Phelan  
APPLICANT: Melissa A. Stavrosnik  
APPLICANT: James A. Welis  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
TITLE OF INVENTION: Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/965,056  
FILING DATE: 05-No. 6271198-1997  
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

US-08-965-056-16

Query Match 64.2%; Score 290; DB 4; Length 268;  
Best Local Similarity 98.1%; Pred. No. 4.4e-22;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKWSLWMPNI 52  
|||||  
Db 154 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKWSLWMPNI 205

RESULT 4  
US-09-272-342B-7  
Sequence 7, Application US/09272342B  
Patent No. 6294341  
GENERAL INFORMATION:  
APPLICANT: YU, YEON-GYU

APPLICANT: KIM, SUNG-HOU  
APPLICANT: YU, JAE-RYEON  
TITLE OF INVENTION: METHOD FOR DETECTING A SUBSTANCE HAVING AN ACTIVITY TO  
TITLE OF INVENTION: INHIBIT HIV INFECTION USING IMMUNOASSAY AND VARIANT  
TITLE OF INVENTION: PROTEIN USED FOR SAID METHOD  
FILE REFERENCE: 2901-0125-0  
CURRENT APPLICATION NUMBER: US/09/272,342B  
CURRENT FILING DATE: 1999-03-19  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 7  
LENGTH: 345  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-272-342B-7

Query Match 64.2%; Score 290; DB 4; Length 345;  
Best Local Similarity 98.1%; Pred. No. 5.8e-22;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKWSLWMPNI 52  
|||||  
Db 113 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKWSLWMPNI 164

RESULT 5  
US-08-589-446-8  
Sequence 8, Application US/08589446  
Patent No. 5614413  
GENERAL INFORMATION:

APPLICANT: Morrow, Casey D.  
TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC  
TITLE OF INVENTION: ACID AND METHODS OF MAKING AND  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/589,446  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/087,009

ATTORNEY/AGENT INFORMATION:  
NAME: Geary III, William C.  
REGISTRATION NUMBER: 31,359  
REFERENCE/DOCKET NUMBER: UAG-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: Linear

MOLECULE TYPE: protein  
US-08-589-446-8

Query Match 64.2%; Score 290; DB 1; Length 519;  
Best Local Similarity 98.1%; Pred. No. 9.3e-22;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



QY 1 NHTTMEWDREINNTSLIHSLIEEQNQEKNEQELLELDKWSALMNWNI 52  
 ||||:||||||||||||||||||||||||||||||||||||||||||  
 Db 420 NHTTMEWDREINNTSLIHSLIEEQNQEKNEQELLELDKWSALMNWNI 477

RESULT 6  
US-08-444-882-8  
; Sequence 8, Application US/08444882  
; Patent No. 5622705

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; 00100

```

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,8823  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/087,0099  
INVENTOR: JAMES H. HARRIS

NAME: Gearty III, William C.  
REGISTRATION NUMBER: 31,359  
REFERENCE/DOCKET NUMBER: UAG-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-444-882-8

```

Query Match      64.2%; Score
Best local Similarity 98.1%; Predicted
Matches 51; Conservative 1; M

QY      1 NHTTWEMDREINNTSLIHSLEESON
        |||||:|||||:|||||:|||||:
Db      420 NHTTWEMDREINNTSLIHSLEESON

RESULT 7
US-08-389-459A-8
; Sequence 8; Application US/08389459A
; Patent No. 5817512

```

APPLICANT: Morrow, Casey D. and Porter, Donna, C.  
TITLE OF INVENTION: ENCAPSULATED RECOMBINANT POLIOVIRUS  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
TITLE OF INVENTION: USING SAME  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MASSACHUSETTS

```

1      COUNTRY:  USA
2      ZIP:  02109
3
4      COMPUTER READABLE FORM:
5      MEDium TYPE:  floppy disk
6      COMpuTER:  IBM PC compatible
7      OPERATING SYSTEM:  PC-DOS/MS-DOS
8      SOFTWARE:  ASCII
9
10     CURRENT APPLICATION DATA:
11     APPLICATION NUMBER:  US/08/389,459A
12     FILING DATE:  15-FEB-1995
13     CLASSIFICATION:  435
14     PRIOR APPLICATION DATA:
15     APPLICATION NUMBER:  US 08/087,009
16     FILING DATE:  01-JUL-1993
17     CLASSIFICATION:  435
18
19     ATTORNEY/AGENT INFORMATION:
20     NAME:  Silverl, Jean M.
21     REGISTRATION NUMBER:  39,030
22     REFERENCE/DOCKET NUMBER:  UAG-004C
23     TELECOMMUNICATION INFORMATION:
24     TELEPHONE:  (617) 227-7400
25     TELEFAX:  (617) 227-5941
26
27     INFORMATION FOR SEQ ID NO:  8:
28
29     SEQUENCE CHARACTERISTICS:
30     LENGTH:  519 amino acids
31     TYPE:  amino acid
32     TOPOLOGY:  linear
33
34     MOLECULE TYPE:  protein
35
36     US-08-389-459A-8

```

Query Match	64.2%	Score 290	DB 2	Length 519
Best Local Similarity	98.1%	Pred. No.	9.3e-22	
Matches	51	Conservative	1	Mismatches 0
				Indels 0
				Gaps 0

Db 420 NHTTWMENDREINNYTSLIHSIESONOCERKEDELLEDKASLWNNINI 47  
 RESULT 8  
 US-08-987-867A-8  
 Sequence 8, Application US/08987867A  
 Patent No. 6063384  
 GENERAL INFORMATION:  
 APPLICANT: C. Morrow et al.  
 TITLE OF INVENTION: ENCAPSITATED RECOMBINANT VIRAL  
 TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
 TITLE OF INVENTION: USING SAME  
 NUMBER OF SEQUENCES: 23  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: LAHIVE & COCKFIELD  
 STREET: 28 STATE STREET  
 CITY: BOSTON  
 STATE: MASSACHUSETTS  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/987,867A  
 FILING DATE: 09-DEC-1997  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/087,009  
 FILING DATE: 01-JUL-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Myers, Louis  
 REGISTRATION NUMBER: 35,965  
 REFERENCE/DOCKET NUMBER: UAP-004CPDV  
 TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-987-867A-8

Query Match 64.2%; Score 290; DB 3; Length 519;  
Best Local Similarity 98.1%; Pred. No. 9.3e-22;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESQNOQEKNEQELLELDKMSLWMWFI 52  
Db 420 NHTTWLEMDREINNYTSLIHSLEESQNOQEKNEQELLELDKMSLWMWFI 471

RESULT 9  
US-07-916-098A-2  
Sequence 2, Application US/07916098A  
Patent No. 5871732  
GENERAL INFORMATION:  
APPLICANT: BURKLY, LINDA C.  
APPLICANT: CHISHOLM, PATRICIA L.  
APPLICANT: THOMAS, DAVID W.  
APPLICANT: ROSA, MARGARET D.  
APPLICANT: ROSA, JOSEPH J.  
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN  
TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.  
STREET: 10 SOUTH WACKER DRIVE  
CITY: CHICAGO  
STATE: ILLINOIS  
COUNTRY: U.S.A.  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/916, 098A  
FILING DATE: July 24, 1992  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08843  
FILING DATE: No. 5871732ember 27, 1991  
CLASSIFICATION: 424  
APPLICATION NUMBER: 07/618,542  
FILING DATE: No. 5871732ember 27, 1990  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: JOHN J. MC DONNELL  
REGISTRATION NUMBER: 26,949  
REFERENCE/DOCKET NUMBER: 92,310-G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 715-1000  
TELEFAX: (312) 715-1234  
TELEX: 910/221-5317  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 856 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-916-098A-2

Query Match 64.2%; Score 290; DB 2; Length 856;  
Best Local Similarity 98.1%; Pred. No. 1.7e-21;

Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESQNOQEKNEQELLELDKMSLWMWFI 52  
Db 624 NHTTWLEMDREINNYTSLIHSLEESQNOQEKNEQELLELDKMSLWMWFI 675

RESULT 10  
US-09-337-387-11  
Sequence 11, Application US/09337387  
Patent No. 6420545  
GENERAL INFORMATION:  
APPLICANT: HOXIE, James A.  
APPLICANT: LABRANCHE, Celia C.  
APPLICANT: DOMS, Robert W.  
APPLICANT: HOFFMAN, Trevor L.  
TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
TITLE OF INVENTION: THERAPEUTICS  
FILE REFERENCE: Hoxie 9596-104U1 (0282)  
CURRENT APPLICATION NUMBER: US/09/337,387  
PRIOR FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: US 09/317,556  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 11  
LENGTH: 856  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-337-387-11

Query Match 64.2%; Score 290; DB 4; Length 856;  
Best Local Similarity 98.1%; Pred. No. 1.7e-21;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESQNOQEKNEQELLELDKMSLWMWFI 52  
Db 624 NHTTWLEMDREINNYTSLIHSLEESQNOQEKNEQELLELDKMSLWMWFI 675

RESULT 11  
US-08-472-240A-10  
Sequence 10, Application US/08472240A  
Patent No. 6284248  
GENERAL INFORMATION:  
APPLICANT: KIENY, Marie-Paule  
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
TITLE OF INVENTION: 9P160 VARIANT  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,240A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/956,483  
FILING DATE: 31-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 017753-055  
TELECOMMUNICATION INFORMATION:



Thu Mar 27 11:37:57 2003

us-09-877-606-4.raii

Page 6

**OY**    1 NHTTWLEWDREINNTSLIHSLEEÖNQOEKNEÖELLELDKWSLWNNFNI 52  
       | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :  
**Db**    85 NNMTWEWDREINNTSLIHSLEESÖNQOEKNEÖELLELDKWSLWNNFNI 136

Search completed: March 18, 2003, 14:50:05  
Job time : 21.5172 secs

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:36:16 ; Search time 19.6798 Seconds  
(without alignments)  
199.079 Million cell updates/sec

Title: US-09-877-606-4

Perfect score: 452  
Sequence: 1 NHHTWLEMDREINNYTSLIH.....SKYIHENIARIKILIGER 85

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications\_Ai.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	64.2	268	10	US-09-854-816-16
2	290	64.2	519	10	US-09-756-551A-8
3	290	64.2	856	10	US-09-476-242-1
4	277	61.3	56	10	US-09-779-451-4
5	277	61.3	268	10	US-09-854-816-17
6	277	61.3	268	10	US-09-854-816-18
7	277	61.3	345	10	US-09-779-451-8
8	277	61.3	853	9	US-10-003-035-33
9	277	61.3	1101	9	US-10-003-035-53
10	277	61.3	1186	9	US-10-003-035-55
11	274	60.6	268	10	US-09-854-816-19
12	260	57.5	268	10	US-09-854-816-13
13	260	57.5	269	10	US-09-854-816-46
14	256	56.6	46	10	US-09-779-451-41
15	255	56.4	269	10	US-09-854-816-43
16	254	56.2	268	10	US-09-854-816-9
17	253	56.0	269	10	US-09-854-816-12
18	252	55.8	268	10	US-09-854-816-26
19	252	55.8	269	10	US-09-854-816-28

20	252	55.8	847	10	US-09-476-242-2	Sequence 2, Appl
21	251	55.5	233	10	US-09-854-816-50	Sequence 50, Appl
22	251	55.5	619	10	US-09-891-609-4	Sequence 4, Appl
23	251	55.5	646	10	US-09-891-609-2	Sequence 4, Appl
24	248	54.9	269	10	US-09-854-816-44	Sequence 8, Appl
25	247	54.6	268	10	US-09-854-816-8	Sequence 41, Appl
26	247	54.6	268	10	US-09-854-816-41	Sequence 42, Appl
27	246	54.4	269	10	US-09-854-816-42	Sequence 32, Appl
28	245	54.2	269	10	US-09-854-816-32	Sequence 34, Appl
29	245	54.2	269	10	US-09-854-816-34	Sequence 38, Appl
30	244	54.0	267	10	US-09-854-816-38	Sequence 15, Appl
31	243	53.8	267	10	US-09-854-816-15	Sequence 45, Appl
32	243	53.8	269	10	US-09-854-816-45	Sequence 6, Appl
33	242	53.5	269	10	US-09-854-816-6	Sequence 11, Appl
34	241	53.3	267	10	US-09-854-816-11	Sequence 35, Appl
35	241	53.3	268	10	US-09-854-816-35	Sequence 20, Appl
36	241	53.3	268	10	US-09-854-816-20	Sequence 72, Appl
37	240	53.1	268	10	US-09-854-816-72	Sequence 10, Appl
38	240	53.1	269	10	US-09-854-816-33	Sequence 49, Appl
39	239	52.9	269	10	US-09-854-816-10	Sequence 40, Appl
40	238	52.7	269	10	US-09-854-816-31	Sequence 21, Appl
41	237	52.4	233	10	US-09-854-816-49	Sequence 22, Appl
42	237	52.4	268	10	US-09-854-816-40	Sequence 25, Appl
43	237	52.4	269	10	US-09-854-816-21	
44	237	52.4	269	10	US-09-854-816-22	
45	237	52.4	269	10	US-09-854-816-25	

## ALIGNMENTS

RESULT 1  
US-09-854-816-16  
; Sequence 16, Application US/09854816  
; Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovashnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 MB floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9681  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:



J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-854-816-17  
Query Match 61.3%; Score 277; DB 10; Length 268;  
Best Local Similarity 94.2%; Pred. No. 1.5e-19;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NHTTWLEMDREINNYTSLHSLIEESONOEKNEQELLELDKWSLMMWNI 52  
I: ||:|||||  
Db 154 NNMTWMDREINNYTSLHSLIEESONOEKNEQELLELDKWSLMMWNI 205  
RESULT 6  
US-09-854-816-18  
Sequence 18, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-854-816-18  
Query Match 61.3%; Score 277; DB 10; Length 268;  
Best Local Similarity 94.2%; Pred. No. 1.5e-19;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NHTTWLEMDREINNYTSLHSLIEESONOEKNEQELLELDKWSLMMWNI 52  
I: ||:|||||  
Db 154 NNMTWMDREINNYTSLHSLIEESONOEKNEQELLELDKWSLMMWNI 205  
RESULT 7  
US-09-779-451-8  
Sequence 8, Application US/09779451  
Patent No. US20020094521A1  
GENERAL INFORMATION:  
APPLICANT: Wild, Carl T.  
APPLICANT: Allaway, Graham P.  
TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors  
FILE REFERENCE: 1900.0300003  
CURRENT APPLICATION NUMBER: US/09/779,451  
FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: US 60/235,901  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/181,543  
PRIOR FILING DATE: 2000-02-10  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 8  
LENGTH: 345  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-779-451-8  
Query Match 61.3%; Score 277; DB 10; Length 345;  
Best Local Similarity 94.2%; Pred. No. 1.9e-19;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NHTTWLEMDREINNYTSLHSLIEESONOEKNEQELLELDKWSLMMWNI 52  
I: ||:|||||  
Db 113 NNMTWMDREINNYTSLHSLIEESONOEKNEQELLELDKWSLMMWNI 164  
RESULT 8  
US-10-003-035-33  
Sequence 33, Application US/10003035  
Patent No. US2002015127A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Danher

```
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 33
; LENGTH: 853
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env with multi-clade V3 loops
US-10-003-035-33

Query Match          61.3%; Score 277; DB 9; Length 853;
Best Local Similarity 94.2%; Pred. No. 5.2e-19;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLSIESONQOEKNEQELLELDKWSLWNI 52
   1:|||||
Db 773 NMTWMEWDREINNYTSLIHSLSIESONQOEKNEQELLELDKWSLWNI 824

RESULT 9
US-10-003-035-53
; Sequence 53, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53
; LENGTH: 1101
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Tat
US-10-003-035-53

Query Match          61.3%; Score 277; DB 9; Length 1101;
Best Local Similarity 94.2%; Pred. No. 6.9e-19;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLSIESONQOEKNEQELLELDKWSLWNI 52
   1:|||||
Db 986 NMTWMEWDREINNYTSLIHSLSIESONQOEKNEQELLELDKWSLWNI 1037

RESULT 10
US-10-003-035-55
; Sequence 55, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
```

```
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 55
; LENGTH: 1186
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Tat/Rev
US-10-003-035-55

Query Match          61.3%; Score 277; DB 9; Length 1186;
Best Local Similarity 94.2%; Pred. No. 7.5e-19;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLSIESONQOEKNEQELLELDKWSLWNI 52
   1:|||||
Db 986 NMTWMEWDREINNYTSLIHSLSIESONQOEKNEQELLELDKWSLWNI 1037

RESULT 11
US-09-854-816-19
; Sequence 19, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; J. Kevin Judice
; Robert S. McDowell
; J. Christopher Phelan
; Melissa A. Starovasnik
; James A. Wells
; Making Same
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, Ph.D., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 19:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 268 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-854-816-19

Query Match          60.6%; Score 274; DB 10; Length 268;
Best Local Similarity 92.3%; Pred. No. 2.8e-19;
Matches 48; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNYTSLIHSLSIESONQOEKNEQELLELDKWSLWNI 52
```



Db 154 NNHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFI 205

RESULT 12  
US-09-854-816-13

; Sequence 13, Application US/09854816  
; Patent No. US20020151473A1  
; GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California

COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.

REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:

LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 13:  
US-09-854-816-13

Query Match 57.5%; Score 260; DB 10; Length 268;  
Best Local Similarity 60.0%; Pred. No. 6.3e-18;

Matches 51; Conservative 11; Mismatches 11; Indels 12; Gaps 1;

Db 154 NNHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFI 203

QY 1 NHHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFIKIKIIEOK 60

Db 154 NNHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFI 203

QY 61 IBEILSKIVHENEIARIKRLIGER 85

Db 204 --DTWKWIKIFIMIVGSLIGLR 226

RESULT 13  
US-09-854-816-46

; Sequence 46, Application US/09854816  
; Patent No. US20020151473A1  
; GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell

Query Match 56.6%; Score 256; DB 10; Length 46;

J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California

COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001

CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.

REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:

LENGTH: 269 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 46:  
US-09-854-816-46

Query Match 57.5%; Score 260; DB 10; Length 269;  
Best Local Similarity 88.5%; Pred. No. 6.3e-18;

Matches 46; Conservative 5; Mismatches 1; Indels 0; Gaps 0;

Db 155 NNHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFI 206

QY 1 NHHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFI 52

Db 155 NNHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFI 206

RESULT 14  
US-09-779-451-41

; Sequence 41, Application US/09779451  
; Patent No. US20020094521A1  
; GENERAL INFORMATION:

APPLICANT: Wild, Carl T.  
APPLICANT: Allway, Graham P.

TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors  
FILE REFERENCE: 1900.0300003

CURRENT APPLICATION NUMBER: US/09/779,451  
CURRENT FILING DATE: 2001-08-17

PRIOR APPLICATION NUMBER: US 60/235,901  
PRIOR FILING DATE: 2000-09-28

PRIOR APPLICATION NUMBER: US 60/181,543  
PRIOR FILING DATE: 2000-02-10

NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.0

SEQ ID NO 41  
LENGTH: 46  
TYPE: PRT

ORGANISM: Human immunodeficiency virus type 1  
US-09-779-451-41

Query Match 56.6%; Score 256; DB 10; Length 46;

Db 155 NNHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFI 206

QY 1 NHHTWLEWREINNTSLHSLIDESONOEKNEQELLELDKWSLMMWFI 52

Best Local Similarity 97.8%; Pred. No. 2.2e-18;  
Matches 45; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 5 WLEMDREINNTSLIHSILIESONQOEKNEQELLELDKWSLWNI 50  
|:|||||  
Db 1 WMEMDREINNTSLIHSILIESONQOEKNEQELLELDKWSLWNI 46

## RESULT 15

US-09-854-816-43  
; Sequence 43, Application US/09854816  
; Patent No. US20020151473A1

## GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same

NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

## COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/854,816

FILING DATE: 15-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, PhD., Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P1005R2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 269 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 43:

US-09-854-816-43

QY 1 NHTTWLEMDREINNTSLIHSILIESONQOEKNEQELLELDKWSLWNI 52  
|:|||||  
Db 155 NNMWTWMEEREIDNYTGLIYSLIESONQOEKNEQELLELDKWSLWNI 206

Search completed: March 18, 2003, 14:52:57  
Job time : 20.6798 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:35:13 ; Search time 3964.85 Seconds  
(without alignments)  
623.917 Million cell updates/sec

Title: US-09-877-606-4  
Perfect score: 452  
Sequence: 1 NHTTWLEMDREINNTSLIH.....SKIYHIENIARIKKLIGER 85

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DEV=xlh  
-O=/cgn2.1/USPTO/US09877606/runat\_18032003\_134800\_17655/app.query.fasta\_1.1059  
-DB=genEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-ONITS=bits -START=1 -END=-1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=Pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptlo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09877606 @CGN 1.1 3852 @runat\_18032003\_134800\_17655 -NCPu=6 -ICPU=3  
-NO\_XLPRX -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: genEmbl:\*  
2: gb\_ba:\*  
3: gb\_hcg:\*  
4: gb\_in:\*  
5: gb\_com:\*  
6: gb\_ov:\*  
7: gb\_pat:\*  
8: gb\_ph:\*  
9: gb\_pl:\*  
10: gb\_pr:\*  
11: gb\_ro:\*  
12: gb\_sts:\*  
13: gb\_sy:\*  
14: gb\_un:\*  
15: gb\_vl:\*  
16: em\_ba:\*  
17: em\_fun:\*  
18: em\_hum:\*  
19: em\_in:\*  
20: em\_com:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pat:\*  
24: em\_ph:\*  
25: em\_dl:\*  
26: em\_to:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_hcg\_hum:\*  
31: em\_hcg\_inv:\*  
32: em\_hcg\_other:\*  
33: em\_hcg\_mus:\*  
34: em\_hcg\_pln:\*  
35: em\_hcg\_rod:\*  
36: em\_hcg\_mam:\*  
37: em\_hcg\_vrt:\*  
38: em\_sy:\*  
39: em\_hcgo\_hum:\*  
40: em\_hcgo\_mus:\*  
41: em\_hcgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	64.2	1098	9	HUMIG2GP M18404 Human IgG2
2	290	64.2	1260	6	AR000534 AR000534 Sequence
3	290	64.2	1568	6	AR044676 AR044676 Sequence
4	290	64.2	1568	6	I38637 I38637 Sequence
5	290	64.2	1568	6	I40608 I40608 Sequence
6	290	64.2	1568	6	AR035189 AR035189 Sequence
7	290	64.2	2571	14	AF358143 AF358143 HIV-1 clo
8	290	64.2	2571	14	AF358144 AF358144 HIV-1 clo
9	290	64.2	2573	14	AF358140 AF358140 HIV-1 clo
10	290	64.2	2573	14	AF358142 AF358142 HIV-1 clo
11	290	64.2	2680	14	HIV1012032 HIV1012032 Human Immun
12	290	64.2	2681	14	HIV1012033 HIV1012033 Human Immun
13	290	64.2	2691	14	HIV1012030 HIV1012030 Human Immun
14	290	64.2	2692	14	HIV1012031 HIV1012031 Human Immun
15	290	64.2	2694	6	I83400 I83400 Sequence 1
16	290	64.2	3084	6	I83401 I83401 Sequence 2
17	290	64.2	3156	6	E01088 E01088 Nucleic aci
18	290	64.2	3156	6	HIVHXB3 M14100 Human Immun
19	290	64.2	6031	12	XXU19867 U19867 Cloning vec
20	290	64.2	6505	6	AX427796 AX427796 Sequence
21	290	64.2	6505	6	AX427777 AX427777 Sequence
22	290	64.2	6505	6	AX427778 AX427778 Sequence
23	290	64.2	6532	6	AX427831 AX427831 Sequence
24	290	64.2	6538	6	AX427792 AX427792 Sequence
25	290	64.2	6538	6	AX427795 AX427795 Sequence
26	290	64.2	6541	6	AX427793 AX427793 Sequence
27	290	64.2	6577	6	AX427794 AX427794 Sequence
28	290	64.2	6610	6	AX427828 AX427828 Sequence
29	290	64.2	6610	6	AX427829 AX427829 Sequence
30	290	64.2	6613	6	AX427830 AX427830 Sequence
31	290	64.2	6616	6	AX427789 AX427789 Sequence
32	290	64.2	6619	6	AX427786 AX427786 Sequence
33	290	64.2	6619	6	AX427790 AX427790 Sequence
34	290	64.2	6634	6	AX427791 AX427791 Sequence
35	290	64.2	6646	6	AX427799 AX427799 Sequence
36	290	64.2	6649	6	AX427827 AX427827 Sequence
37	290	64.2	6655	6	AX427787 AX427787 Sequence
38	290	64.2	6658	6	AX427788 AX427788 Sequence
39	290	64.2	6670	6	AX427780 AX427780 Sequence
40	290	64.2	6688	6	AX427824 AX427824 Sequence
41	290	64.2	6691	6	AX427821 AX427821 Sequence
42	290	64.2	6691	6	AX427825 AX427825 Sequence
43	290	64.2	6697	6	AX427783 AX427783 Sequence
44	290	64.2	6706	6	AX427826 AX427826 Sequence
45	290	64.2	6709	6	AX427781 AX427781 Sequence

RESULT 1

## ALIGNMENTS

HUMIG26P  
LOCUS HUMIG26P 1098 bp DNA linear PRI 08-NOV-1994  
DEFINITION Human IgG2 lambda antibody (Ib8.env reactive) gp41 coding region  
DNA.  
ACCESSION M18404  
VERSION M18404.1 GI:184839  
KEYWORDS antibody; glycoprotein.  
SOURCE Human (HIV seropositive male) DNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.  
REFERENCE  
AUTHORS Banapour, B., Rosenthal, K., Rabin, L., Sharma, V., Young, L.,  
Fernandez, J., Engelman, E., McGrath, M., Reyes, G. and Lifson, J.  
TITLE Characterization and epitope mapping of a human monoclonal antibody  
reactive with the envelope glycoprotein of human immunodeficiency  
virus  
JOURNAL J. Immunol. 139 (12), 4027-4033 (1987)  
MEDLINE 88088754  
PUBMED 2447158  
FEATURES  
source Location/Qualifiers  
1..1098  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/map="14932.33"  
1..387  
/gene="IGHG2"  
1..387  
/gene="IGHG2"  
mat\_peptide /note="gp41; G00-119-338"  
BASE COUNT 337 a 209 c 295 g 257 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3.52e-22 Length: 1098  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 9 Gaps: 0  
US-09-877-606-4 (1-85) x HUMIG26P (1-1098)  
QY 1 AsnHsThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrHsrSerLeuIleHis 20  
|||||  
Db 337 AATCAGACGACCTGATGGAGTGGACAGAGAAATTAACAAATTCACAAAGCTTAATACAC 396  
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnIleLysAsnGlnGlnIleLeuLeu 40  
|||||  
Db 397 TCCTTAATTGAAGATGCAAAACACAGCAAGAAATGACAAAGATTAATTGGAATTA 456  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 457 GATTAATGGCAAGTTGTGGAATTTGTTAACAATA 492  
RESULT 2  
LOCUS AR000534 1260 bp DNA linear PAT 04-DEC-1998  
DEFINITION Sequence 7 from patent US 5736391.  
ACCESSION AR000534  
VERSION AR000534.1 GI:3963065  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1260)  
AUTHORS Essex, M.E., Yu, X. and Lee, T.-H.  
TITLE HIV gp41 mutants  
JOURNAL Patent: US 5736391-A 7 07-APR-1998;  
FEATURES Location/Qualifiers  
1..1260  
source /organism="unknown"  
BASE COUNT 398 a 214 c 357 g 291 t

ORIGIN  
Alignment Scores:  
Pred. No.: 4.05e-22 Length: 1260  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 6 Gaps: 0  
US-09-877-606-4 (1-85) x AR000534 (1-1260)  
QY 1 AsnHsThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrHsrSerLeuIleHis 20  
|||||  
Db 464 AATCAGACGACCTGATGGAGTGGACAGAGAAATTAACAAATTCACAAAGCTTAATACAC 523  
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnIleLysAsnGlnGlnIleLeuLeu 40  
|||||  
Db 524 TCCTTAATTGAAGATGCAAAACACAGCAAGAAATGACAAAGATTAATTGGAATTA 583  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 584 GATTAATGGCAAGTTGTGGAATTTGTTAACAATA 619  
RESULT 3  
LOCUS AR044676 1568 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 7 from patent US 5817512.  
ACCESSION AR044676  
VERSION AR044676.1 GI:5966141  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Morrow, C.D., Porter, D.C. and Ansardi, D.C.  
TITLE Encapsidated recombinant viral nucleic acid and methods of making  
and using same  
JOURNAL Patent: US 5817512-A 7 06-OCT-1998;  
FEATURES Location/Qualifiers  
1..1568  
source /organism="unknown"  
BASE COUNT 563 a 267 c 357 g 381 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5.06e-22 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 6 Gaps: 0  
US-09-877-606-4 (1-85) x AR044676 (1-1568)  
QY 1 AsnHsThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrHsrSerLeuIleHis 20  
|||||  
Db 1264 AATCAGACGACCTGATGGAGTGGACAGAGAAATTAACAAATTCACAAAGCTTAATACAC 1323  
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnIleLysAsnGlnGlnIleLeuLeu 40  
|||||  
Db 1324 TCCTTAATTGAAGATGCAAAACACAGCAAGAAATGACAAAGATTAATTGGAATTA 1383  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATTAATGGCAAGTTGTGGAATTTGTTAACAATA 1419  
RESULT 4  
LOCUS I38637 1568 bp DNA linear PAT 13-MAY-1997  
DEFINITION Sequence 7 from patent US 5614413.  
ACCESSION I38637  
VERSION I38637.1 GI:2084691

KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Morrow, C.D.  
TITLE Encapsidated recombinant poliovirus nucleic acid and methods of making and using same  
JOURNAL Patent: US 5614413-A 7 25-MAR-1997;  
FEATURES Location/Qualifiers  
source 1..1568  
/organism="unknown"  
BASE COUNT 563 a 267 c 357 g 381 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.06e-22 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
Gaps: 0  
DB:

US-09-877-606-4 (1-85) x 138637 (1-1568)

QY 1 AsnHstHrThrPleuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
|||||  
Db 1264 AATCAGACAGCAGCTGGATGGAGTGGACAGAGAAATTACACATTAATACAC 1323

QY 21 SerLeuIleGluIleUserGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeu 40  
|||||  
Db 1324 TCCTTAATTGAAGATCGCAAAACAGCAAGAAATGACAGAAATATTGGAATTA 1383

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATAAATGGCGAGTTGTGGAAATGGTTTAACATA 1419

RESULT 5  
LOCUS 140608 1568 bp DNA linear PART 13-MAY-1997  
DEFINITION Sequence 7 from patent US 5622705.  
ACCESSION 140608  
VERSION 140608.1 GI:2082088  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Morrow, C.D.  
TITLE Encapsidated recombinant poliovirus nucleic acid and methods of making and using same  
JOURNAL Patent: US 5622705-A 7 22-APR-1997;  
FEATURES Location/Qualifiers  
source 1..1568  
/organism="unknown"  
BASE COUNT 563 a 267 c 357 g 381 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.06e-22 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
Gaps: 0  
DB:

US-09-877-606-4 (1-85) x 140608 (1-1568)

QY 1 AsnHstHrThrPleuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
|||||  
Db 1264 AATCAGACAGCAGCTGGATGGAGTGGACAGAGAAATTACACATTAATACAC 1323

QY 21 SerLeuIleGluIleUserGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeu 40  
|||||  
Db 1324 TCCTTAATTGAAGATCGCAAAACAGCAAGAAATGACAGAAATATTGGAATTA 1383

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATAAATGGCGAGTTGTGGAAATGGTTTAACATA 1419

RESULT 5  
LOCUS 140608 1568 bp DNA linear PART 13-MAY-1997  
DEFINITION Sequence 7 from patent US 5622705.  
ACCESSION 140608  
VERSION 140608.1 GI:2082088  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1568)  
AUTHORS Morrow, C.D.  
TITLE Encapsidated recombinant poliovirus nucleic acid and methods of making and using same  
JOURNAL Patent: US 5622705-A 7 22-APR-1997;  
FEATURES Location/Qualifiers  
source 1..1568  
/organism="unknown"  
BASE COUNT 563 a 267 c 357 g 381 t  
ORIGIN

Alignment Scores:  
Pred. No.: 5.06e-22 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
Gaps: 0  
DB:

US-09-877-606-4 (1-85) x 140608 (1-1568)

Db 1324 TCCTTAATTGAAGATCGCAAAACAGCAAGAAATGACAGAAATTAATGGAATTA 1383

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1384 GATAAATGGCGAGTTGTGGAAATGGTTTAACATA 1419

RESULT 6  
LOCUS AR035189 2571 bp DNA linear PART 29-SEP-1999  
DEFINITION Sequence 1 from patent US 5871732.  
ACCESSION AR035189  
VERSION AR035189.1 GI:5951857  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2571)  
AUTHORS Burkly, L.C., Chisholm, P.L., Thomas, D.W., Rosa, M.D. and Rosa, J.J.  
TITLE Anti-Cd4 antibody homologs useful in prophylaxis and treatment of AIDS, ARC and HIV infection  
JOURNAL Patent: US 5871732-A 1 16-FEB-1999;  
FEATURES Location/Qualifiers  
source 1..2571  
/organism="unknown"

BASE COUNT 892 a 439 c 616 g 624 t  
ORIGIN

Alignment Scores:  
Pred. No.: 8.36e-22 Length: 2571  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
Gaps: 0  
DB:

US-09-877-606-4 (1-85) x AR035189 (1-2571)

QY 1 AsnHstHrThrPleuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
|||||  
Db 1870 AATCAGACAGCAGCTGGATGGAGTGGACAGAGAAATTACACATTAATACAC 1929

QY 21 SerLeuIleGluIleUserGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeu 40  
|||||  
Db 1930 TCCTTAATTGAAGATCGCAAAACAGCAAGAAATGACAGAAATATTGGAATTA 1989

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 1990 GATAAATGGCGAGTTGTGGAAATGGTTTAACATA 2025

RESULT 7  
LOCUS AF358143 2571 bp DNA linear VRL 03-MAY-2001  
DEFINITION HIV-1 clone C48.1 from France envelope glycoprotein (env) pseudogene, complete sequence.  
ACCESSION AF358143  
VERSION AF358143.1 GI:13936875  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1.  
REFERENCE 1 (bases 1 to 2571)  
AUTHORS Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate lentivirus group.  
TITLE Attman-Onal, Y., Cheynet, V. and Verrier, B.  
Mutations and transcriptional alterations associated with the downregulation of HIV-1 envelope glycoprotein expression following acute cytopathic effects  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2571)  
AUTHORS Attman-Onal, Y., Cheynet, V. and Verrier, B.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2001) UMR 2142 CNRS/Biomerieux, CERVI-IFR4, 21 Avenue Tony Garnier, Lyon, Rhone 69365, France

FEATURES  
Source Location/Qualifiers  
1..2571  
/organism="Human immunodeficiency virus type 1"  
/strain="HXB2"  
/db\_xref="taxon:11676"  
/clone="C48.1"  
/cell\_line="HL60"  
/country="France"  
1..309  
/gene="env"  
/note="envelope glycoprotein; contains premature stop codons"  
Bases: 891 a 438 c 619 g 623 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.36e-22 Length: 2571  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 14 Gaps: 0  
US-09-877-606-4 (1-85) x AF358143 (1-2571)  
QY 1 AsnHsthrThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrHrSerLeuIleHis 20  
|||||  
DB 1870 AATCACACGACCTGGATGGAGTGGACAGACAATTAACATTAACACAGCTTAAACAC 1929  
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluIlysAsnGluGlnGluLeuGluLeu 40  
|||||  
DB 1930 TCCTTAATGGAAGATCGCAAAACACGACAGAAAAGAAATGAACAATTAATGGAATTA 1989  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
DB 1990 GATAAATGGCAAGTTCTGCAATTGCTTAAACATA 2025  
RESULT 8 2571 bp DNA linear VRL 03-MAY-2001  
AF358144  
LOCUS HIV-1 clone C48.2 from France envelope glycoprotein (env)  
DEFINITION pseudogene, complete sequence.  
ACCESSION AF358144  
VERSION AF358144.1 GI:13936876  
KEYWORDS Human immunodeficiency virus type 1.  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Viruses: Retroviridae; Lentivirus; Primate  
Lentivirus group.  
REFERENCE 1 (bases 1 to 2571)  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Mutations and transcriptional alterations associated with the downregulation of HIV-1 envelope glycoprotein expression following acute cytopathic effects  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2571)  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2001) UMR 2142 CNRS/BioMerieux, CERVI-IFR74, 21 Avenue Tony Garnier, Lyon, Rhone 69365, France  
FEATURES  
Source Location/Qualifiers  
1..2571  
/organism="Human immunodeficiency virus type 1"  
/strain="HXB2"  
/db\_xref="taxon:11676"  
/clone="C48.2"  
/cell\_line="HL60"  
/country="France"  
1..309  
/gene="env"  
/note="envelope glycoprotein; contains premature stop codons"

BASE COUNT 894 a 436 c 616 g 625 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.36e-22 Length: 2571  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 14 Gaps: 0  
US-09-877-606-4 (1-85) x AF358144 (1-2571)  
QY 1 AsnHsthrThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrrHrSerLeuIleHis 20  
|||||  
DB 1870 AATCACACGACCTGGATGGAGTGGACAGACAATTAACATTAACACAGCTTAAACAC 1929  
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluIlysAsnGluGlnGluLeuGluLeu 40  
|||||  
DB 1930 TCCTTAATGGAAGATCGCAAAACACGACAGAAAAGAAATGAACAATTAATGGAATTA 1989  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
DB 1990 GATAAATGGCAAGTTCTGCAATTGCTTAAACATA 2025  
RESULT 9 2573 bp DNA linear VRL 03-MAY-2001  
AF358140  
LOCUS HIV-1 clone C8.1 from France envelope glycoprotein (env)  
DEFINITION pseudogene, complete sequence.  
ACCESSION AF358140  
VERSION AF358140.1 GI:13936871  
KEYWORDS Human immunodeficiency virus type 1.  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Viruses: Retroviridae; Lentivirus; Primate  
Lentivirus group.  
REFERENCE 1 (bases 1 to 2573)  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Mutations and transcriptional alterations associated with the downregulation of HIV-1 envelope glycoprotein expression following acute cytopathic effects  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2573)  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Direct Submission  
JOURNAL Submitted (09-MAR-2001) UMR 2142 CNRS/BioMerieux, CERVI-IFR74, 21 Avenue Tony Garnier, Lyon, Rhone 69365, France  
FEATURES  
Source Location/Qualifiers  
1..2573  
/organism="Human immunodeficiency virus type 1"  
/strain="HXB2"  
/db\_xref="taxon:11676"  
/clone="C8.1"  
/cell\_line="HL60"  
/country="France"  
1..2570  
/gene="env"  
/note="envelope glycoprotein; contains premature stop codons"  
BASE COUNT 894 a 442 c 614 g 623 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.37e-22 Length: 2573  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-4 (1-85) x AF358140 (1-2573)

QY 1 AsnHstHtThrTrrpLeuGltTrrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
|||||  
Db 1870 AATCACACGACCTGGATGGAGTGAGACAGAAATTAAACAACTTAACAGCTTAATACAC 1929

QY 21 SerLeuIleGluGlnSerGlnAsnGlnGlnGluLysAsnGlnGlnGlnLeuLeuGluLeu 40  
|||||  
Db 1930 TCCTTAATGGAAGATCGCAAAACGACAGAAAGAAATGACAAATATTATTGGAATTA 1989

QY 41 AspLysTrrpAlaSerLeuTrrpAsnTrrpPheAsnIle 52  
|||||  
Db 1990 GATAAATGGCAAGTTTGTGAATTGTTTAACATA 2025

RESULT 10  
AF358142 2573 bp DNA linear VRL 03-MAY-2001

LOCUS  
DEFINITION HIV-1 clone C8.3 from France envelope glycoprotein (env)

ACCESSION  
AF358142  
pseudogene, complete sequence.

VERSION  
AF358142.1 GI:13936874

KEYWORDS

SOURCE  
ORGANISM human immunodeficiency virus type 1.  
Human immunodeficiency virus type 1  
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
1 (bases 1 to 2573)

REFERENCE  
AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Mutations and transcriptional alterations associated with the downregulation of HIV-1 envelope glycoprotein expression following acute cytopathic effects

JOURNAL  
REFERENCE Unpublished  
2 (bases 1 to 2573)

AUTHORS Ataman-Onal,Y., Cheynet,V. and Verrier,B.  
TITLE Direct Submission  
SUBMITTED (09-MAR-2001) UMR 2142 CNRS/Biomerieux, CERV1-IFR74, 21 Avenue Tony Garnier, Lyon, Rhone 69365, France

FEATURES  
LOCATION/Qualifiers  
source 1..2573  
/organism="Human immunodeficiency virus type 1"  
/strain="HXB2"  
/db\_xref="taxon:11676"  
/clone="C8.3"  
/cell\_line="HL60"  
/country="France"  
1..2570  
/gene="env"  
/note="envelope glycoprotein: contains premature stop codons"  
/pseudo

BASE COUNT 891 a 440 c 617 g 625 t

ORIGIN

Alignment Scores:  
Pred. No.: 8.37e-22 Length: 2573  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-4 (1-85) x AF358142 (1-2573)

QY 1 AsnHstHtThrTrrpLeuGltTrrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
|||||  
Db 1870 AATCACACGACCTGGATGGAGTGAGACAGAAATTAAACAACTTAATACAC 1929

QY 21 SerLeuIleGluGlnSerGlnAsnGlnGlnGluLysAsnGlnGlnGlnLeuLeuGluLeu 40  
|||||  
Db 1930 TCCTTAATGGAAGATCGCAAAACGACAGAAAGAAATGACAAATATTATTGGAATTA 1989

QY 41 AspLysTrrpAlaSerLeuTrrpAsnTrrpPheAsnIle 52  
|||||

Db 1990 GATAAATGGCAAGTTTGTGAATTGTTTAACATA 2025

RESULT 11  
HIV1U12032 2680 bp ss-RNA linear VRL 28-SEP-1995

LOCUS  
DEFINITION Human immunodeficiency virus type 1 isolate LM881 from infected lab worker, envelope glycoprotein (env) gene, partial cds.

ACCESSION  
U12032  
U12032.1 GI:665490

VERSION

KEYWORDS

SOURCE  
ORGANISM Human immunodeficiency virus type 1.  
Human immunodeficiency virus type 1  
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
1 (bases 1 to 2680)

REFERENCE  
AUTHORS Reltz,M., Hall,L., Robert-Guroff,M., Lautenberger,J., Hahn,B., Shaw,G., Kong,L., Weiss,S., Waters,D., Gallo,R. and Blattner,W.  
TITLE Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HTLV type IIIB)

JOURNAL  
MEDLINE AIDS Res. Hum. Retroviruses 10 (9), 1143-1155 (1994)  
PUBMED 95127297  
7826699

REFERENCE  
AUTHORS 2 (bases 1 to 2680)  
Muller,K.E.  
TITLE Direct Submission  
SUBMITTED (06-JUN-1994) Katherine E. Muller, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA

COMMENT  
FEATURES  
source 1..2680  
/organism="Human immunodeficiency virus type 1"  
/isolate="LM881"  
/db\_xref="taxon:11676"  
437..2680  
/gene="env"  
437..2680  
/gene="env"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AA1668.1"  
/db\_xref="GI:665491"  
/translation="MRVKEKYOLRRMGWRMGTMGLMILMCSATEKIMLVTVYGVV  
WKEATTLFCASDAKAYDTEVNWATACVTDNDPEVYVNTENMMKNMVE  
OMHEDTILMDSQSLKPCVKLIPCYSLCTDLKNFTNNSGGMIMKGEIKNSFN  
FTDNVKTIIIVLQNSVEJNCKPNNNTGRRIIOGPRTFTYIGKIGNMROHONS  
RAKNVNTLKQJASKLREQYGNKTIIFQSSGDEIVTHSFCGFEFFCNSTOLF  
STMFSTGSNNTSGSDITLPCRIKOIINMOEYKAMAYAPISQIRCSSVITGLL  
TRDGSNNNGSEIFPRGGDMRDMRSELYKKYKIBPLGVAPRKARVYOREKRA  
VIGALPLGFLGACAGSTMGASMTLTGAROLISGIVNOONLLPRTAOCHLOTV  
WGIKOLQRIITLAVERYLKDQOLGITMGCSGLICTTAYPMNMSMSKSLERIMNTTW  
MEWREINNTYSLHSLIEESONDERKEDELELDKMASLMNWNITNMLWYVILFI  
MIVGVLGRLIVFAVLSTVNRVROGSPLSQTHLPFGDPBPGIEEGEERDRDR  
SIRLVNCS"

BASE COUNT 965 a 451 c 623 g 641 t

ORIGIN

Alignment Scores:  
Pred. No.: 8.72e-22 Length: 2680  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-4 (1-85) x HIV1U12032 (1-2680)

QY 1 AsnHstHtThrTrrpLeuGltTrrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
|||||  
Db 2294 AATCACACGACCTGGATGGAGTGAGACAGAAATTAAACAACTTAATACAC 2353

QY 21 Serleu11eglugluserGlnAsnglncglulysasnnglucgluleu 40  
LOCUS HIVU12033 2681 bp ss-RNA linear VRL 28-SEP-1995  
DEFINITION Human immunodeficiency virus type 1 isolate LM88-2 from infected lab worker, envelope glycoprotein (env) pseudogene, partial cds.  
ACCESSION U12033.1 GI:665492  
VERSION 012033  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1.  
VIRUSES; Retrovirdae; Lentiviridae; Lentivirus; Primate  
REFERENCE  
AUTHORS 1 (bases 1 to 2681)  
Reitz,M., Hall,L., Robert-Guroff,M., Lautenberger,J., Hahn,B., Shaw,G., Kong,L., Weiss,S., Waters,D., Gallo,R. and Blattner,W. Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIB)  
JOURNAL AIDS Res. Hum. Retroviruses 10 (9), 1143-1155 (1994)  
MEDLINE 95127297  
PUBMED 7826699  
REFERENCE 2 (bases 1 to 2681)  
AUTHORS Mulder,K.E.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-1994) Katherine E. Mulder, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
COMMENT Envelope sequence, lambda clone from Sali to BamHI.  
FEATURES  
source location/Qualifiers  
1..2681  
/organism="Human immunodeficiency virus type 1"  
/isolate="LM88-2"  
/db\_xref="taxon:11676"  
437..2681  
/gene="env"  
437..>2681  
/gene="env"  
/pseudo  
/codon\_start=1  
/product="envelope glycoprotein"  
BASE COUNT 966 a 451 c 622 g 642 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 8.73e-22 Length: 2681  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 14 Gaps: 0  
US-09-877-606-4 (1-85) x HIVU12033 (1-2681)  
QY 1 Asnh1sthrtrtprleuglutrpsaparglullesasnnytrtrserleu1lehis 20  
LOCUS HIVU12033 2691 bp ss-RNA linear VRL 28-SEP-1995  
DEFINITION Human immunodeficiency virus type 1 isolate LM851 from infected lab worker, envelope glycoprotein (env) gene, partial cds.  
ACCESSION U12033.1 GI:665486  
VERSION 012033  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1.  
VIRUSES; Retrovirdae; Lentiviridae; Lentivirus; Primate  
REFERENCE  
AUTHORS 1 (bases 1 to 2691)  
Reitz,M., Hall,L., Robert-Guroff,M., Lautenberger,J., Hahn,B., Shaw,G., Kong,L., Weiss,S., Waters,D., Gallo,R. and Blattner,W. Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIB)  
JOURNAL AIDS Res. Hum. Retroviruses 10 (9), 1143-1155 (1994)  
MEDLINE 95127297  
PUBMED 7826699  
REFERENCE 2 (bases 1 to 2691)  
AUTHORS Mulder,K.E.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-1994) Katherine E. Mulder, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
COMMENT Envelope sequence, lambda clone from Sali to BamHI.  
FEATURES  
source location/Qualifiers  
1..2691  
/organism="Human immunodeficiency virus type 1"  
/isolate="LM851"  
/db\_xref="taxon:11676"  
436..2691  
/gene="env"  
436..>2691  
/gene="env"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AA17666.1"  
/db\_xref="GI:665487"  
/translation="MRVKEKYQHLMRWGMRGTMGLMIGCATGKLTWTVYGVV  
WKETATTLTFCASDAKAVDEVHVVWVTHACVPDPDPEVVLVNVENFMKNDWE  
OMEDIISLMDOSLAKPCVKLTPLCYSLKCDLNDNTNSSGRMIMEKEIINCSFN  
ISRSIRGKQKEVAFPKLIDITPINDTSTYTLTSCNTSVYACRKYSPEDPHYC  
APAGPALIKNNKTPNGCTGCTVSVYQCHGIRPVYRSTLLNGSLAEERYIRSN  
FTDNATTVIVOLNOSTEINCTRPNNNTKIRIRORGFVFTIGKGNRQHCNIS  
RAKWNNTLKOIASLREDFGNNTTIFKSSGDDPELVTHSFNCGEFCYCNSTOLEN  
STWFNSWTSEGSNTEGSDPTITLPCRIKOIIMMOEVKAMYPAPISGIRCSNIT  
GLLTRDGGNNNGSEIFRPGGDMRNMSELYKKYVKEIPLGVAPTAKRRVOR  
EKRAVIGALFELFGLAAGSTMGASMTLTVORLISGIVOOONMLRAIEAOHL  
QLTWVGKIQOARILAVERYLKRQOQLGTCGSKLICTTAVPMNSMSKSLDEQTN  
HTTWEMDRFINNTVSLTSLIESONQDKNEQLELELDKRNKSLMMWENTIMWLTI  
KIFIMVGLVGLRIVFAVLSTVNRKQGSPLSFYTHLPTPGGPPREGIEEGGER  
DRDRSIRLVNGS"

QY 21 Serleu11eglugluserGlnAsnglncglulysasnnglucgluleu 40  
LOCUS HIVU12033 2691 bp ss-RNA linear VRL 28-SEP-1995  
DEFINITION Human immunodeficiency virus type 1 isolate LM851 from infected lab worker, envelope glycoprotein (env) gene, partial cds.  
ACCESSION U12033.1 GI:665486  
VERSION 012033  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1.  
VIRUSES; Retrovirdae; Lentiviridae; Lentivirus; Primate  
REFERENCE  
AUTHORS 1 (bases 1 to 2691)  
Reitz,M., Hall,L., Robert-Guroff,M., Lautenberger,J., Hahn,B., Shaw,G., Kong,L., Weiss,S., Waters,D., Gallo,R. and Blattner,W. Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIB)  
JOURNAL AIDS Res. Hum. Retroviruses 10 (9), 1143-1155 (1994)  
MEDLINE 95127297  
PUBMED 7826699  
REFERENCE 2 (bases 1 to 2691)  
AUTHORS Mulder,K.E.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-1994) Katherine E. Mulder, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
COMMENT Envelope sequence, lambda clone from Sali to BamHI.  
FEATURES  
source location/Qualifiers  
1..2691  
/organism="Human immunodeficiency virus type 1"  
/isolate="LM851"  
/db\_xref="taxon:11676"  
436..2691  
/gene="env"  
436..>2691  
/gene="env"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AA17666.1"  
/db\_xref="GI:665487"  
/translation="MRVKEKYQHLMRWGMRGTMGLMIGCATGKLTWTVYGVV  
WKETATTLTFCASDAKAVDEVHVVWVTHACVPDPDPEVVLVNVENFMKNDWE  
OMEDIISLMDOSLAKPCVKLTPLCYSLKCDLNDNTNSSGRMIMEKEIINCSFN  
ISRSIRGKQKEVAFPKLIDITPINDTSTYTLTSCNTSVYACRKYSPEDPHYC  
APAGPALIKNNKTPNGCTGCTVSVYQCHGIRPVYRSTLLNGSLAEERYIRSN  
FTDNATTVIVOLNOSTEINCTRPNNNTKIRIRORGFVFTIGKGNRQHCNIS  
RAKWNNTLKOIASLREDFGNNTTIFKSSGDDPELVTHSFNCGEFCYCNSTOLEN  
STWFNSWTSEGSNTEGSDPTITLPCRIKOIIMMOEVKAMYPAPISGIRCSNIT  
GLLTRDGGNNNGSEIFRPGGDMRNMSELYKKYVKEIPLGVAPTAKRRVOR  
EKRAVIGALFELFGLAAGSTMGASMTLTVORLISGIVOOONMLRAIEAOHL  
QLTWVGKIQOARILAVERYLKRQOQLGTCGSKLICTTAVPMNSMSKSLDEQTN  
HTTWEMDRFINNTVSLTSLIESONQDKNEQLELELDKRNKSLMMWENTIMWLTI  
KIFIMVGLVGLRIVFAVLSTVNRKQGSPLSFYTHLPTPGGPPREGIEEGGER  
DRDRSIRLVNGS"



QY 41 AsplystrPalaserLeutPAsnTrpPheasn1le 52  
Db 2425 GATTAATGGCAGAGTTGTGGAATGTTAACAATA 2460  
RESULT 14  
LOCUS HIV1U12031  
DEFINITION Human immunodeficiency virus type 1 isolate LM852 from infected lab worker, envelope glycoprotein (env) gene, partial cds.  
ACCESSION U12031.1 GI:665488  
VERSION 1  
KEYWORDS Human immunodeficiency virus type 1.  
SOURCE Human immunodeficiency virus type 1  
ORGANISM Viruses: Retroid viruses; Retroviridae; Lentivirus; primate  
REFERENCE 1 (bases 1 to 2692)  
AUTHORS Reitz,M., Hall,L., Robert-Guroff,M., Lautenberger,J., Hahn,B., Shaw,G., Kong,L., Weiss,S., Waters,D., Gallo,R. and Blattner,W.  
TITLE Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIIB)  
JOURNAL AIDS Res. Hum. Retroviruses 10 (9), 1143-1155 (1994)  
MEDLINE 95127297  
PUBMED 7826699  
REFERENCE 2 (bases 1 to 2692)  
AUTHORS Mulder K.E.  
TITLE Direct Submission  
JOURNAL Submitted (06-JUL-1994) Katherine E. Mulder, Los Alamos National Laboratory, HIV Sequence Database, T-10, Mail Stop K710, Los Alamos, NM 87501, USA  
COMMENT Envelope sequence, lambda clone from SalI to BamHI.  
FEATURES  
source 1..2692  
Location/Qualifiers  
1..2692  
/organism="Human immunodeficiency virus type 1"  
/isolate="LM852"  
/db\_xref="taxon:11676"  
437..2692  
/gene="env"  
437..2692  
/gene="env"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AA076667.1"  
/db\_xref="GI:665488"  
/translation="MRVKEVYOHIMRWGMRGTMILGIMICSAPEKLMVTVYGVY  
WKEATITLFCASDAKAYDREVENNATHNVCYTPDRPOVAVLYNTENTNKNMAYE  
QNHEDIISLMDOSLKPVCVLTPLCYLCTDLDKNDNTSSSGRMIMERKGEIKNSFN  
ISTSLRGVQKRYAFYKLDIIPIDNDITSYLTSCNTSVITQACPKVSEPIPIHYC  
APAGFAILCKNNKTFNGTGPCTNVSTVQCTHGIRPVSTOLLNGLAEVEVIRSAN  
FIDNAKTIIVOLNOSIEINCTRPNNTRRIRIORGPAFTVIGLIGMROAHNIS  
RAKNNNTLKQIASKLREORGNNKTIIFKSSGDPRIYVHSNGGEPFISQSTOLF  
STWPNSTWTEGSSNTEGSDITTLPRKQIITNMDEVGKAMTAPISQICSSNIT  
GLLTRDGGNNNGSEIFRPGGDMKMDNRSELYRKVKIIEPLGVAAPRKARVVOR  
EKRAVIGALFLGFLGAGSTWGAASMTLVQAROLSGIVQOONMLLAIRAQHL  
QLTVWGIKQLARILAVERYLKDQDLIGMCGSKLICCTAVPMNANSGNKLEQIWN  
HTTMEEMDEINNTSLHSLIEESONOEKMEDELLEDKWASLMMNENINMLMYI  
KTFIMIVGSLGRLIVFAVLSTVNNVROGYSPLSFQTHLPTGGDPDRPGIEEGGER  
DDRSIRLVNGS"

BASE COUNT 970 a 457 c 624 g 641 t  
ORIGIN

Alignment Scores:  
Pred. No.: 8,76e-22 Length: 2692  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-4 (1-85) x HIV1U12031 (1-2692)

QY 1 AsnHISThrTrtPLeuGIuTrPAspARgGluIleAsnAsnTyTrtHrSerLeuIleHIS 20

Db 2306 AATCACACGACCTCGAAGAGGTGGACAGAAATTAAACACTTAATACAC 2365  
QY 21 SerLeuIleGluIuSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeu 40  
Db 2366 TCCTTAATTGAAGATGCAAAACCAAGAAAGATGACAAAGATTATTGGAATTA 2425  
QY 41 AsplystrPalaserLeutPAsnTrpPheasn1le 52  
Db 2426 GATTAATGGCAGAGTTGTGGAATGTTAACAATA 2461  
RESULT 15  
LOCUS 183400 2694 bp DNA linear PAT 10-AUG-1998  
DEFINITION Sequence 1 from patent US 5714316.  
ACCESSION 183400  
VERSION 183400.1 GI:3406930  
KEYWORDS Unknown.  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2694)  
AUTHORS Weiner,D., Williams,W. and Levy,D.N.  
TITLE Chimeric envelope proteins for viral targeting  
JOURNAL Patent: US 5714316-A 1 03-FEB-1998:  
FEATURES  
source 1..2694  
Location/Qualifiers  
1..2694  
/organism="unknown"  
BASE COUNT 889 a 515 c 657 g 633 t  
ORIGIN

Alignment Scores:  
Pred. No.: 8,77e-22 Length: 2694  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 6 Gaps: 0

US-09-877-606-4 (1-85) x 183400 (1-2694)

QY 1 AsnHISThrTrtPLeuGIuTrPAspARgGluIleAsnAsnTyTrtHrSerLeuIleHIS 20  
Db 1978 AATCACACGACCTCGAAGAGGTGGACAGAAATTAAACACTTAATACAC 2037  
QY 21 SerLeuIleGluIuSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeu 40  
Db 2038 TCCTTAATTGAAGATGCAAAACCAAGAAAGATGACAAAGATTATTGGAATTA 2097  
QY 41 AsplystrPalaserLeutPAsnTrpPheasn1le 52  
Db 2098 GATTAATGGCAGAGTTGTGGAATGTTAACAATA 2133

Search completed: March 25, 2003, 15:49:35  
Job time : 3966.85 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 11:26:42 ; Search time 331.626 Seconds  
(without alignments)  
577.217 Million cell updates/sec

Title: US-09-877-606-4  
Perfect score: 452  
Sequence: 1 NHTWLEMDREINNTSLIH.....SKYHIENIARIKKLIGER 85

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2.1/USPTO-SPOOL/US09877606/runatc\_18032003\_134800\_17645/app\_query.fasta.1.1059  
-DB=N\_Geneseq\_101002 -QFMT=fastap -SUFFIX=ing -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09877606.ecgn.1.1.564\_etrunatc\_18032003\_134800\_17645 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -IONCLOG -DEV\_TIMEROUT=120  
-NARN\_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-XGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query			DB ID	Description
		Match	Length	DB		
1	290	64.2	1260	15	AA066275	Sequence of HIV-1
2	290	64.2	1568	16	AA080575	HIV-1 env gene. H
3	290	64.2	1568	17	AAT33295	HIV-1 env gene. H
4	290	64.2	2694	19	AAV09364	GMSF/HIV-1 chimera
5	290	64.2	3084	19	AAV09365	HIV-1/CD4 chimera
6	290	64.2	3156	7	AA60128	Sequence of the en
7	290	64.2	3246	21	AAA48622	Lentivirus vector
8	290	64.2	6460	24	ABK91482	Modified HIV prote
9	290	64.2	6505	24	ABK91463	Modified HIV prote
10	290	64.2	6505	24	ABK91464	Modified HIV prote
11	290	64.2	6532	24	ABK91517	Modified HIV prote
12	290	64.2	6538	24	ABK91478	Modified HIV prote
13	290	64.2	6538	24	ABK91481	Modified HIV prote
14	290	64.2	6541	24	ABK91479	Modified HIV prote
15	290	64.2	6577	24	ABK91480	Modified HIV prote
16	290	64.2	6610	24	ABK91514	Modified HIV prote
17	290	64.2	6610	24	ABK91515	Modified HIV prote
18	290	64.2	6613	24	ABK91516	Modified HIV prote
19	290	64.2	6616	24	ABK91475	Modified HIV prote
20	290	64.2	6619	24	ABK91472	Modified HIV prote
21	290	64.2	6619	24	ABK91476	Modified HIV prote
22	290	64.2	6634	24	ABK91477	Modified HIV prote
23	290	64.2	6646	24	ABK91465	Modified HIV prote
24	290	64.2	6649	24	ABK91513	Modified HIV prote
25	290	64.2	6655	24	ABK91473	Modified HIV prote
26	290	64.2	6658	24	ABK91474	Modified HIV prote
27	290	64.2	6670	24	ABK91466	Modified HIV prote
28	290	64.2	6688	24	ABK91510	Modified HIV prote
29	290	64.2	6691	24	ABK91507	Modified HIV prote
30	290	64.2	6691	24	ABK91511	Modified HIV prote
31	290	64.2	6697	24	ABK91469	Modified HIV prote
32	290	64.2	6706	24	ABK91512	Modified HIV prote
33	290	64.2	6709	24	ABK91467	Modified HIV prote
34	290	64.2	6712	24	ABK91470	Modified HIV prote
35	290	64.2	6715	24	ABK91471	Modified HIV prote
36	290	64.2	6727	24	ABK91508	Modified HIV prote
37	290	64.2	6730	24	ABK91509	Modified HIV prote
38	290	64.2	6736	24	ABK91468	Modified HIV prote
39	290	64.2	6769	24	ABK91504	Modified HIV prote
40	290	64.2	6774	24	ABK91536	Modified HIV prote
41	290	64.2	6784	24	ABK91505	Modified HIV prote
42	290	64.2	6787	24	ABK91455	Modified HIV prote
43	290	64.2	6787	24	ABK91506	Modified HIV prote
44	290	64.2	6808	24	ABK91503	Modified HIV prote
45	290	64.2	6846	24	ABK91537	Modified HIV prote

## ALIGNMENTS

RESULT 1	AA066275	standard. DNA; 1260 BP.
ID	AA066275	
AC	AA066275	
XX		
DT	28-DEC-1994	(first entry)
XX		
DE	Sequence of HIV-1 encoding the gp41 polypeptide.	
XX		
KW	Human Immunodeficiency virus; HIV-1; AIDS; lentivirus; TM protein;	
KW	extracellular protein; transmembrane protein; gp41; ds.	
XX		
OS	HIV-1.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	3..1260
FT		/*tag- a

```
FT      CDS /note= "AAR53783"
FT      1..1260
FT      /*tag= b
FT      /note= "AAR53784"
FT      CDS 2..1269
FT      /*tag= c
FT      /product= gp41
FT      /note= "encoded by bps 128..1160 (AAR53785)"
FT
FT
FT      W09412533-A.
FT
FT      PD 09-JUN-1994.
FT
FT      PE 12-JAN-1993; 93WO-US00212.
FT
FT      PR 23-NOV-1992; 92US-0979975.
FT
FT      PA (HARD ) HARVARD COLLEGE.
FT
FT      PI Essex ME, Lee TH, Yu X;
FT
FT      DR WPI: 1994-200197/24.
FT      P-PSDB; AAR53783, AAR53784, AAR53785.
FT
FT      PT Method for treating HIV patients - comprises administration of
FT      mutated gp41 polypeptide
FT
FT      PS Claim 20; Fig 1; 54pp; English.
FT
FT      XX
CC      The inventors claim a method of treating a patient infected with HIV
CC      by administering a mutated gp41 polypeptide or a therapeutic
CC      composition comprising nucleic acid encoding the mutant gp41
CC      polypeptide in an expressible genetic construction. The mutant gp41
CC      polypeptide contains a deletion of at least one AA in at least one
CC      of the following regions of wild type gp41 (AAR53783): AAs 844-856;
CC      814-856; 796-856; 776-856; 753-856; or 710-856, effective to either
CC      disrupt viral replication or HIV or disrupt the assembly of viral
CC      Env proteins in an HIV infected cell. AAQ6275 corresp. to bps 7631-
CC      8890 of wt HIV-1.
CC
CC      SQ Sequence 1260 BP; 398 A; 214 C; 356 G; 292 T; 0 other;
CC
CC      Alignment Scores:
CC      Pred. No.: 1.54e-25 Length: 1260
CC      Score: 290.00 Matches: 51
CC      Percent Similarity: 100.00% Conservative: 1
CC      Best Local Similarity: 98.08% Mismatches: 0
CC      Query Match: 64.16% Indels: 0
CC      DB: 15 Gaps: 0
CC
US-09-877-606-4 (1-85) x AAQ6275 (1-1260)
QY      1 AsnHsTrHrTrPLeuGluTrPAsPaRgSJuIleAsnAsnYrThrSerLeuIleHis 20
DB      464 AATCAGCAGACCTGATGAGTGGAGCAGAGAAATTAACATTACACAGCTTAATAC 523
QY      21 SerLeuIleGluJusSerGlnAsnGlnGlnIuLysAsnGluGlnGluLeuGluLeu 40
DB      524 TCCTTAATTGAAGATGCAAAACAGCAAGAAAGAAATGAACAATATTGGAATT 583
QY      41 AspLysTrPAlaSerLeuTrPAsnTrPheAsnIle 52
DB      584 GATAAATGGCAAGTTGTGGAATTGTTAACATA 619
RESULT 2
ID      AAQ80575 standard; DNA; 1568 BP.
XX
XX      AAQ80575;
XX
XX      13-SEP-1995 (first entry)
XX
XX      HIV-1 env gene.
DE
```

```
XX      Encapsulation; Human immunodeficiency virus; type 1; immunogenic;
XX      poliovirus; ss.
XX
XX      Human immunodeficiency virus type 1.
XX
XX      Key Location/Qualifiers
XX      CDS 7..1565
XX      /*tag= a
XX
XX      CA2125344-A.
XX
XX      PD 02-JAN-1995.
XX
XX      PE 07-JUN-1994; 94CA-2125344.
XX
XX      PR 01-JUL-1993; 93US-0087009.
XX
XX      PA (UABR-) UAB RES FOUND.
XX
XX      PI Morrow CD;
XX
XX      DR WPI: 1995-099021/14.
XX      P-PSDB; AAR69997.
XX
XX      PT Method for encapsulating recombinant polio:virus nucleic acid -
XX      useful for providing comps. to stimulate immune response.
XX
XX      PS Disclosure; Page 38; 62pp; English.
XX
XX      The sequence is that of the HIV-1 env gene. The gene is used in a
XX      method (claimed) to encapsulate poliovirus cDNA to make it more
XX      immunogenic. The method involves providing a poliovirus DNA
XX      sequence which lacks the fragment encoding the protein necessary for
XX      viral encapsulation and an expression vector lacking an infectious
XX      poliovirus genome but encoding a portion of the encapsulating
XX      protein. The poliovirus DNA also encodes a HIV-1 gene, in this case
XX      the env gene, so the recombinant virus can induce a response to both
XX      poliovirus and HIV.
XX      See also AAQ80573-7.
XX
XX      SQ Sequence 1568 BP; 563 A; 267 C; 357 G; 381 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 2.03e-25 Length: 1568
XX      Score: 290.00 Matches: 51
XX      Percent Similarity: 100.00% Conservative: 1
XX      Best Local Similarity: 98.08% Mismatches: 0
XX      Query Match: 64.16% Indels: 0
XX      DB: 16 Gaps: 0
XX
US-09-877-606-4 (1-85) x AAQ80575 (1-1568)
QY      1 AsnHsTrHrTrPLeuGluTrPAsPaRgSJuIleAsnAsnYrThrSerLeuIleHis 20
DB      1264 AATCAGCAGACCTGATGAGTGGAGCAGAGAAATTAACATTACACAGCTTAATAC 1323
QY      21 SerLeuIleGluJusSerGlnAsnGlnGlnIuLysAsnGluGlnGluLeuGluLeu 40
DB      1324 TCCTTAATTGAAGATGCAAAACAGCAAGAAAGAAATGAACAATATTGGAATT 1383
QY      41 AspLysTrPAlaSerLeuTrPAsnTrPheAsnIle 52
DB      1384 GATAAATGGCAAGTTGTGGAATTGTTAACATA 1419
RESULT 3
ID      AAT33295 standard; cDNA; 1568 BP.
XX
XX      AAT33295;
XX
XX      12-NOV-1996 (first entry)
XX
```

DE HIV-1 env gene.  
XX  
XX Poliovirus; encapsidation: capsid; vaccine; genetic immunisation;  
KM HIV-1; human immunodeficiency virus type 1; env gene; ss.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
FH Key Location/Qualifiers  
FT CDS 7..1563  
FT /tag= a  
XX  
XX WO9625173-A1.  
XX  
XX 22-AUG-1996.  
XX  
XX 13-FEB-1996; 96WO-US01895.  
XX  
XX 15-FEB-1995; 95US-0389459.  
XX  
XX (UABR-) UAB RES FOUND.  
XX  
XX Anardi DC, Morrow CD, Porter DC;  
XX  
XX WPI: 1996-393136/39.  
XX P-PSDB: AAM00181.  
XX  
XX Encapsidation of recombinant polio:virus nucleic acid for use in  
PT vaccines - using a polio:virus nucleic acid which lacks the p1  
PT capsid region and an expression system which provides the region  
XX  
XX  
XX Disclosure: Page 61-63; 108pp; English.  
XX  
XX  
XX cDNA sequences (AAT33293-95) respectively code for the gag, pol and  
CC env proteins (AAM00129-81) of HIV-1. They can be used to substitute  
CC the p1 capsid gene of poliovirus in recombinant poliovirus nucleic  
CC acids (rPNAs). Such rPNAs are encapsidated by introduction into  
CC a host cell together with a vaccinia virus or plasmid vector  
CC encoding the poliovirus p1 capsid precursor protein. Encapsidated  
CC rPNAs are useful for genetic immunisation, stimulating an immune  
CC response to the HIV-1 protein.  
XX  
XX  
SQ Sequence 1568 BP; 563 A; 267 C; 357 G; 381 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 2,03e-25 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 17 Gaps: 0  
XX  
US-09-877-606-4 (1-85) x AAT33295 (1-1568)  
  
QY 1 AsnHstHtHtRpleuGluTrpAspArgGluIleAsnAsnTyThrSerLeuIleHis 20  
DB 1264 AATTCACACGACCTGGATGGAGTGCGACAGAAATTAACAATTAACAAGCTTAATACAC 1323  
QY 21 SerLeuIleGluGluSerLeuAsnGlnGlnIuLyAsnGlnGlnIuLeuLeuGluLeu 40  
DB 1324 TCCCTTAATTGAAGAATCGCAAAACGCAAGAAAGAAAGAACAAATTAATTGGAATTA 1383  
QY 41 AspLySTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
DB 1384 GATTAATGGCAAGTTTGGAATTGCTTAACATA 1419  
RESULT 4  
AAV09364  
ID AAV09364 standard; DNA: 2694 BP.  
AC AAV09364;  
XX  
XX 13-MAY-1998 (first entry)  
XX

DE GMCSF/HIV-1 chimeric nucleotide sequence.  
XX  
XX Human immunodeficiency virus; HIV; CD4; GMCSF; chimeric; delivery;  
KM granulocyte-macrophage colony stimulating factor; env protein; ss.  
XX  
XX Chimeric - Human immunodeficiency virus type 1.  
OS Chimeric - Homo sapiens.  
XX  
XX US5714316-A.  
XX  
XX 03-FEB-1998.  
XX  
XX 04-NOV-1993; 93US-0147890.  
XX  
XX 04-NOV-1993; 93US-0147890.  
XX  
XX 21-JUN-1991; 91US-0718537.  
XX  
XX (UYPB-) UNIV PENNSYLVANIA.  
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
XX  
XX Levy DN, Weiner D, Williams W;  
XX  
XX WPI: 1998-129852/12.  
XX  
XX Chimeric constructs for HIV proteins - having HIV binding site to  
PT CD4 replaced by ligand, useful for delivery of therapeutic or  
PT diagnostic agents  
XX  
XX Claim 4; Fig 5A-B; 29pp; English.  
XX  
XX This is a chimeric GMCSF/HIV-1 nucleotide construct which comprises an  
CC HIV gp160 envelope gene from which the sequence encoding amino acids  
CC 392-446 or a fragment corresponding to the HIV binding site to CD4 has  
CC been deleted and into which has been inserted a heterologous sequence  
CC encoding a ligand selected from granulocyte-macrophage colony stimulating  
CC factor (GMCSF). A recombinant viral particle can be produced by a host  
CC cell transformed with a vector comprising this chimeric construct and a  
CC regulatory sequence which directs expression of the chimeric construct.  
CC The viral particle is incapable of targeting its native ligand receptor  
CC as the chimeric protein lacks a complete env protein due to the deletion  
CC of the native retroviral binding region. The recombinant viral particle  
CC is non-pathogenic and incapable of recombination in vivo into a  
CC pathogenic form because it lacks the portion of the env protein  
CC responsible for targeting native retroviral cellular receptors and lacks  
CC viral nucleic acids. The viral particles can be used for the delivery of  
CC biological mediators such as therapeutic agents, gene products,  
CC diagnostic labels or toxic agents. They can be used for delivery to, e.g.  
CC pluripotent stem cells (using GMCSF ligand) or the gp120 protein of HIV  
CC envelope (CD4 ligand).  
XX  
XX  
SQ Sequence 2694 BP; 889 A; 515 C; 657 G; 633 T; 0 other;  
  
Alignment Scores:  
Pred. No.: 3.98e-25 Length: 2694  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 19 Gaps: 0  
XX  
US-09-877-606-4 (1-85) x AAV09364 (1-2694)  
  
QY 1 AsnHstHtHtRpleuGluTrpAspArgGluIleAsnAsnTyThrSerLeuIleHis 20  
DB 1978 AATTCACACGACCTGGATGGAGTGCGACAGAAATTAACAATTAACAAGCTTAATACAC 2037  
QY 21 SerLeuIleGluGluSerLeuAsnGlnGlnIuLyAsnGlnGlnIuLeuLeuGluLeu 40  
DB 2038 TCCCTTAATTGAAGAATCGCAAAACGCAAGAAAGAAAGAACAAATTAATTGGAATTA 2097  
QY 41 AspLySTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
DB 2098 GATTAATGGCAAGTTTGGAATTGCTTAACATA 2133

RESULT 5  
AAV09365  
ID AAV09365 standard; DNA: 3084 BP.  
XX  
AC AAV09365;  
XX  
DT 13-MAY-1998 (first entry)  
XX  
DE HIV-1/CD4 chimeric nucleotide sequence.  
XX  
KW Human immunodeficiency virus; HIV; CD4; GMCSF; chimeric; delivery;  
XX granulocyte-macrophage colony stimulating factor; env protein; ss.  
XX  
OS Chimeric - Human immunodeficiency virus type 1.  
OS Chimeric - Homo sapiens.  
XX  
PN US5714316-A.  
XX  
PD 03-FEB-1998.  
XX  
PF 04-NOV-1993; 93US-0147890.  
XX  
PR 04-NOV-1993; 93US-0147890.  
XX 21-JUN-1991; 91US-0718537.  
XX  
PA (UYPE-) UNIV PENNSYLVANIA.  
XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.  
XX  
PI Levy DN, Weiner D, Williams W;  
XX  
DR WPI; 1998-129852/12.  
XX  
PT Chimeric constructs for HIV proteins - having HIV binding site to  
PT CD4 replaced by ligand, useful for delivery of therapeutic or  
PT diagnostic agents  
XX  
PS Claim 5; Fig 6A-C; 29pp; English.  
XX  
CC This is a chimeric HIV-1/CD4 nucleotide construct which comprises an  
CC HIV gp160 envelope gene from which the sequence encoding amino acids  
CC 392-446 or a fragment corresponding to the HIV binding site to CD4 has  
CC been deleted and into which has been inserted a heterologous sequence  
CC encoding a ligand selected from CD4. The ligand can also be selected  
CC from the granulocyte-macrophage colony stimulating factor (GMCSF). A  
CC recombinant viral particle can be produced by a host cell transfected  
CC with a vector comprising this chimeric construct and a regulatory  
CC sequence which directs expression of the chimeric construct. The viral  
CC particle is incapable of targeting its native ligand receptor as the  
CC chimeric protein lacks a complete env protein due to the deletion of the  
CC native retroviral binding region. The recombinant viral particle is  
CC non-pathogenic and incapable of recombination in vivo into a pathogenic  
CC form because it lacks the portion of the env protein responsible for  
CC targeting native retroviral cellular receptors and lacks viral nucleic  
CC acids. The viral particles can be used for the delivery of biological  
CC mediators such as therapeutic agents, gene products, diagnostic labels  
CC or toxic agents. They can be used for delivery to, e.g. pluripotent stem  
CC cells (using GMCSF ligand) or the gp120 protein of HIV envelope (CD4  
CC ligand).  
XX  
SQ Sequence 3084 BP; 1058 A; 559 C; 748 G; 719 T; 0 other;

Alignment Scores:  
Pred. No.: 4.72e-25 Length: 3084  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 19 Gaps: 0

US-09-877-606-4 (1-85) x AAV09365 (1-3084)

QY 1 AsnHisThrThrIleuGluTTPAsPArgGluIleAsnAsnTyrrThrSerLeuIleHis 20

|||||  
Db 2368 AATCACAGACCTGATGAGTGGACAGAAATTAACATTACACAGCTTAAATCAC 2427  
QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuGluLeu 40  
Db 2428 TCCTTAATTGAAGATGCAAAACACGACAGAAAGATGACAGAAATTTGGAATTA 2487  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
Db 2488 GATAAATGGCAACTTGTGTGAATGTTAACAATA 2523

RESULT 6  
AAN60128  
ID AAN60128 standard; DNA: 3156 BP.  
XX  
AC AAN60128;  
XX  
DT 26-JUN-1991 (first entry)  
XX  
DE Sequence of the envelope gene of the HTLV-III proviral genome  
DE (HXB-3).  
XX  
KW AIDS; HIV; LAV; HTLV-III; vaccine; antibody; epitope; antigen;  
XX diagnosis; ss.  
XX  
OS HTLV-III.  
XX  
FH Key Location/Qualifiers  
FT CDS 478..3048  
FT /\*tag= a  
XX  
PN EP199301-A.  
XX  
PD 29-OCT-1986.  
XX  
PF 18-APR-1986; 86EP-0105371.  
XX  
PR 19-APR-1985; 85US-0725021.  
XX  
PA (HOFF) HOFFMANN-LA ROCHE AG.  
PA (USGO) US GOVERNMENT.  
PA (HEAL-) DEPT HEALTH & HUMAN SERV.  
PA (USDH) US DEPT HEALTH & HUMAN.  
XX  
PI Crowl RM, Gallo RC, Reddy EP, Shaw GM, Wong-Staal FY;  
XX  
DR WPI; 1986-286067/44.  
DR P-PSDB; AAP60131.  
XX  
PT Envelope protein of acquired immune deficiency syndrome virus -  
PT useful for improved testing of human blood for antibodies against  
PT virus and as antigen for vaccines  
XX  
PS Claim 35; Fig 1; 46pp; English.  
XX  
CC An expression vector contg. a gene (AAN60128) coding for an envelope  
CC protein of an AIDS virus, and the envelope protein of an AIDS virus  
CC (AAP60131) are claimed. The vector is pref. a member of the pEV/env  
CC family, e.g. pEV1.2 or 3/env 44-640 or 205-640.  
XX  
SQ Sequence 3156 BP; 1095 A; 536 C; 765 G; 760 T; 0 other;

Alignment Scores:  
Pred. No.: 4.85e-25 Length: 3156  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conserves: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 7 Gaps: 0

US-09-877-606-4 (1-85) x AAN60128 (1-3156)

QY 1 AsnHisThrThrIleuGluTTPAsPArgGluIleAsnAsnTyrrThrSerLeuIleHis 20

|||||  
 Db 2347 AATCAGACGCGTGGATGGGTGGACAGAGAAATTACCAATTACACAGCTTAATACAC 2406  
 Qy 21 SerLeu11eGluGluSerGlnAsnGlnGlnGluLysAsnGlnGluLeuLeuGluLeu 40  
 Db 2407 TCCTTAATTGAGAAATCGCAAAACGACAGAAAGAAATGACAGAAATTATTGAAATTA 2466  
 Qy 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
 Db 2467 GATAAATGGGCAAGTTTGTGGAATGCTTAACATA 2502  
 RESULT 7  
 AAA48622  
 ID AAA48622 standard; DNA; 3246 BP.  
 AC AAA48622;  
 XX  
 XX  
 DT 07-DEC-2000 (first entry)  
 XX  
 DE Lentivirus vector pHR' .  
 XX  
 KW Interleukin-4; IL-4; insulin-dependent diabetes mellitus;  
 KW IDDM; systemic lupus erythematosus; rheumatoid arthritis;  
 KW multiple sclerosis; hypercholesterolemia; muscular dystrophy;  
 KW Huntington's disease; Parkinson's disease; immunosuppressive therapy;  
 KW Lentivirus vector; ds.  
 XX  
 OS Human immunodeficiency virus.  
 XX  
 PN WO200034455-A2.  
 PD 15-JUN-2000.  
 XX  
 PF 09-DEC-1999; 99WO-EP09720.  
 XX  
 PR 10-DEC-1998; 98OS-0111678.  
 XX  
 PA (NOVS ) NOVARTIS AG.  
 PA (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.  
 PA (SCRI ) SCRIPS RES INST.  
 XX  
 PI Sarvetnick NE, Gallichan S, Kafri T, Verma I;  
 DR WPI: 2000-423409/36.  
 XX  
 PT Treating diabetes or protecting against insulinitis associated with  
 PT diabetes comprises transducing immunoregulatory genes into pancreatic  
 PT islets via a lentiviral expression vector .  
 XX  
 PS Disclosure; Fig 2; 41pp; English.  
 XX  
 CC The present sequence is the complete sequence of a HIV-based lentivirus  
 CC vector known as pHR' . It can be used, along with the interleukin-4  
 CC (IL-4) coding sequence, for the transduction of pancreatic islet  
 CC cells which can then be grafted into diabetes patients. The presence of  
 CC the IL-4 coding sequence means that IL-4 is produced by the grafted  
 CC cells and they will not experience rejection from the host. Patients who  
 CC are likely to suffer from diabetes in the future can also benefit from  
 CC this treatment. Other diseases which can be treated in a similar manner  
 CC include systemic lupus erythematosus, rheumatoid arthritis, multiple  
 CC sclerosis, hypercholesterolemia, muscular dystrophy, Huntington's  
 CC disease and Parkinson's disease.  
 XX  
 SQ Sequence 3246 BP; 964 A; 665 C; 900 G; 717 T; 0 other;

Alignment Scores:  
 Pred. No.: 5,03e-25 Length: 3246  
 Score: 290.00 Matches: 51  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 98.08% Mismatches: 0  
 Query Match: 64.16% Indels: 0  
 DB: 21 Gaps: 0

US-09-877-606-4 (1-85) x AAA48622 (1-3246)  
 Qy 1 AsnHisThrThrTrpLeuGluTrpAspArgGluIleAsnAsnTrpThrSerLeuIleHis 20  
 Db 1621 AATCAGACGACCTGGATGGGTGGACAGAGAAATTACCAATTACACAGCTTAATACAC 1680  
 Qy 21 SerLeu11eGluGluSerGlnAsnGlnGlnGluLysAsnGlnGluLeuLeuGluLeu 40  
 Db 1681 TCCTTAATTGAGAAATCGCAAAACGACAGAAAGAAATGACAGAAATTATTGAAATTA 1740  
 Qy 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
 Db 1741 GATAAATGGGCAAGTTTGTGGAATGCTTAACATA 1776  
 RESULT 8  
 ABK91482  
 ID ABK91482 standard; DNA; 6460 BP.  
 AC ABK91482;  
 XX  
 XX  
 DT 14-AUG-2002 (first entry)  
 XX  
 DE Modified HIV protein-encoding plasmid DNA #34.  
 XX  
 KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;  
 KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;  
 KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200232943-A2.  
 PD 25-APR-2002.  
 XX  
 PF 14-AUG-2001; 2001WO-US25721.  
 XX  
 PR 14-AUG-2000; 2000US-225097P.  
 PR 14-NOV-2000; 2000US-252115P.  
 PR 28-MAR-2001; 2001US-279257P.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (CHAD/) CHADRAHART B K.  
 XX  
 PI Nabel GJ, Huang Y;  
 DR WPI: 2002-452382/48.  
 XX  
 PT New modified human immunodeficiency virus Env, Gag, Pol or Nef DNA and  
 PT its encoded protein, useful as vaccines for genetic or protein  
 PT immunisation for acquired immunodeficiency syndrome or human  
 PT immunodeficiency virus infection .  
 XX  
 PS Disclosure; Page 451-453; 794pp; English.  
 XX  
 CC The invention relates to a nucleic acid molecule encoding a modified HIV  
 CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
 CC genetic or protein immunisation to a host, respectively. In particular  
 CC these are useful for ameliorating the symptoms of acquired  
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
 CC compositions are useful for treating or preventing HIV infections or  
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding  
 CC plasmid DNA sequences of the invention.  
 XX  
 SQ Sequence 6460 BP; 1600 A; 1742 C; 1656 G; 1462 T; 0 other;

Alignment Scores:  
 Pred. No.: 1,19e-24 Length: 6460  
 Score: 290.00 Matches: 51  
 Percent Similarity: 100.00% Conservative: 1  
 Best Local Similarity: 98.08% Mismatches: 0

Query Match:	64.16%	Indels:	0
DB:	24	Gaps:	0
US-09-877-606-4 (1-85) x ABR91482 (1-6460)			
OY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20			
DB 3261 AACGACACACCTGGATGGAGTGGAGACCGGAGATCAACATACACGACCTGATCCAC 3320			
OY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnLeuLeuGluLeu 40			
DB 3321 AGCCTGATCGAGGAGACCGAACACACAGAGAGACGAGAGAGCTGCTGGAGCTG 3380			
OY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52			
DB 3381 GACAAgTGGCCAGCTGTGGAACTGGTTCAACATC 3416			
RESULT 9			
ID ABR91463 standard; DNA: 6505 BP.			
XX ABR91463:			
XX 14-AUG-2002 (first entry)			
XX Modified HIV protein-encoding plasmid DNA #15.			
XX HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;			
XX Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;			
XX acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.			
XX Synthetic.			
XX WO200232943-A2.			
XX 25-APR-2002.			
XX 14-AUG-2001; 2001WO-US25721.			
XX 14-AUG-2000; 2000US-225097P.			
XX 14-NOV-2000; 2000US-252115P.			
XX 28-MAR-2001; 2001US-279257P.			
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.			
XX (CHAD/) CHADRAABART B K.			
XX Nabel GJ, Huang Y;			
XX WPI: 2002-452382/48.			
XX Disclosure; Page 409-411; 794pp; English.			
XX The invention relates to a nucleic acid molecule encoding a modified HIV			
XX (human immunodeficiency virus) protein. The nucleic acid molecule and its			
XX encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for			
XX genetic or protein immunisation to a host, respectively. In particular			
XX these are useful for ameliorating the symptoms of acquired			
XX immunodeficiency syndrome (AIDS) or HIV infection and generating an			
XX antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,			
XX Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent			
XX compositions are useful for treating or preventing HIV infections or			
XX AIDS. Sequences ABR91449-ABR91624 represent modified HIV protein-encoding			
XX plasmid DNA sequences of the invention.			
XX Sequence 6505 BP; 1638 A; 1775 C; 1638 G; 1454 T; 0 other;			
Alignment Scores:			
1.2e-24			
Length: 6505			

Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	64.16%	Indels:	0
DB:	24	Gaps:	0
US-09-877-606-4 (1-85) x ABR91463 (1-6505)			
QY	1	ASnHstHrTrrTrrPLeuGlUTTPAsPArGlnIlleASnAsNtYrThrSerLeuIleHs	20
Db	3319	AACCCACACACCTGGATGGATGGATGGACCGCAATCAACAACACACGCTGATCCAC	3378
QY	21	SerLeuIleGluGluSerGlnAsnGlnGluGlyAsnGluGlnGluLeuLeuGluLeu	40
Db	3379	AGCCTGATCGAGGAGACCGCAACACACAGCAAGACAGACAGAGAGCTGCTGAGACTG	3438
QY	41	AsPlyStrPaLaSerLeuTrpAsnTrpPheAsnIle	52
Db	3439	GACAAGTGGCCACCTGTGGAACTGGTTCAACATC	3474
RESULT 10			
ID	ABK91464	ABK91464	
XX	ABK91464	standard; DNA; 6505 BP.	
AC	ABK91464;		
DT	14-AUG-2002	(first entry)	
XX			
DE	Modified HIV protein-encoding plasmid DNA #16.		
XX			
KM	HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;		
KM	Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;		
KM	acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.		
XX			
OS	Synthetic.		
PN	WO200232943-A2.		
PD	25-APR-2002.		
XX			
PF	14-AUG-2001; 2001WO-US25721.		
XX			
PR	14-AUG-2000; 2000US-225097P.		
PR	14-NOV-2000; 2000US-252115P.		
PR	28-MAR-2001; 2001US-279257P.		
PA	(USSH ) US DEPT HEALTH & HUMAN SERVICES.		
PA	(CHAD/) CHADRAABARTI B K.		
XX			
PI	Nabel GJ, Huang Y;		
DR	WPI; 2002-452382/48.		
XX			
PT	New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and		
PT	its encoded protein, useful as vaccines for genetic or protein		
PT	immunisation for acquired immunodeficiency syndrome or human		
PT	immunodeficiency virus infection		
XX			
PS	Disclosure; Page 411-413; 794pp; English.		
XX			
CC	The invention relates to a nucleic acid molecule encoding a modified HIV		
CC	(human immunodeficiency virus) protein. The nucleic acid molecule and its		
CC	encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for		
CC	genetic or protein immunisation to a host, respectively. In particular		
CC	these are useful for ameliorating the symptoms of acquired		
CC	immunodeficiency syndrome (AIDS) or HIV infection and generating an		
CC	antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,		
CC	Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent		
CC	compositions are useful for treating or preventing HIV infections or		
CC	AIDS. Sequences ABR91449-ABK91624 represent modified HIV protein-encoding		
CC	plasmid DNA sequences of the invention.		
XX			
XX	Sequence 6505 BP; 1637 A; 1769 C; 1644 G; 1455 T; 0 other;		



Alignment Scores:

Pred. No.:	1.2e-24	Length:	6505
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	64.16%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-4 (1-85) x ABK91464 (1-6505)

OY 1 AsnHstHrThrTrrPleuGlUtrrPaspArgJuIleAsnAsnTyrThrSerLeuIleHis 20  
 DB 3319 AACCAACACACCTGGATGAGTGGAGCCGCCGAGATCAACACTACACCGCTGATCCAC 3378  
 OY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLyAsnGlnGlnGluLeuLeu 40  
 DB 3379 AGCTGATCGAGAGAGCCAGAACCCAGCAGAGAAAGACGACGAGCTGCTGGAGCTG 3438  
 OY 41 AspLySTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
 DB 3439 GACAAAGTGGCCAGCCTGTGTGAACTGTTCAACATC 3474

RESULT 11  
 ABK91517  
 ID ABK91517 standard; DNA: 6532 BP.  
 AC ABK91517;  
 XX 14-AUG-2002 (first entry)  
 DT Modified HIV protein-encoding plasmid DNA #69.  
 DE  
 XX  
 KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;  
 KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;  
 KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.  
 OS Synthetic.  
 XX WO200232943-A2.  
 PN 25-APR-2002.  
 PD 14-AUG-2001; 2001WO-US25721.  
 PF 14-AUG-2000; 2000US-225097P.  
 PR 14-NOV-2000; 2000US-252115P.  
 PR 28-MAR-2001; 2001US-279257P.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (CHAD/) CHADRABARTI B K.  
 XX  
 PI Nabel GJ, Huang Y;  
 DR WPI: 2002-452382/48.  
 XX  
 PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and  
 PT its encoded protein, useful as vaccines for genetic or protein  
 PT immunisation for acquired immunodeficiency syndrome or human  
 PT immunodeficiency virus infection  
 PS Disclosure: Page 536-538; 794pp; English.  
 XX  
 CC The invention relates to a nucleic acid molecule encoding a modified HIV  
 CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
 CC genetic or protein immunisation to a host, respectively. In particular  
 CC these are useful for ameliorating the symptoms of acquired  
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
 CC compositions are useful for treating or preventing HIV infections or  
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding

CC plasmid DNA sequences of the invention.

Sequence	6532 BP; 1610 A; 1762 C; 1679 G; 1481 T; 0 other;
SO	Sequence 6532 BP; 1610 A; 1762 C; 1679 G; 1481 T; 0 other;

Alignment Scores:

Pred. No.:	1.2e-24	Length:	6532
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	64.16%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-4 (1-85) x ABK91517 (1-6532)

OY 1 AsnHstHrThrTrrPleuGlUtrrPaspArgJuIleAsnAsnTyrThrSerLeuIleHis 20  
 DB 3261 AACCAACACACCTGGATGAGTGGAGCCGCCGAGATCAACACTACACCGCTGATCCAC 3320  
 OY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLyAsnGlnGlnGluLeuLeu 40  
 DB 3321 AGCTGATCGAGAGAGCCAGAACCCAGCAGAGAAAGACGACGAGCTGCTGGAGCTG 3380  
 OY 41 AspLySTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
 DB 3381 GACAAAGTGGCCAGCCTGTGTGAACTGTTCAACATC 3416

RESULT 12  
 ABK91478  
 ID ABK91478 standard; DNA: 6538 BP.  
 AC ABK91478;  
 XX 14-AUG-2002 (first entry)  
 DT Modified HIV protein-encoding plasmid DNA #30.  
 DE  
 XX  
 KW HIV; human immunodeficiency virus; gene; ds; circular; cyclic; anti-HIV;  
 KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;  
 KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.  
 OS Synthetic.  
 XX WO200232943-A2.  
 PN 25-APR-2002.  
 PD 14-AUG-2001; 2001WO-US25721.  
 PF 14-AUG-2000; 2000US-225097P.  
 PR 14-NOV-2000; 2000US-252115P.  
 PR 28-MAR-2001; 2001US-279257P.  
 XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
 PA (CHAD/) CHADRABARTI B K.  
 XX  
 PI Nabel GJ, Huang Y;  
 DR WPI: 2002-452382/48.  
 XX  
 PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and  
 PT its encoded protein, useful as vaccines for genetic or protein  
 PT immunisation for acquired immunodeficiency syndrome or human  
 PT immunodeficiency virus infection  
 PS Disclosure: Page 442-444; 794pp; English.  
 XX  
 CC The invention relates to a nucleic acid molecule encoding a modified HIV  
 CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
 CC genetic or protein immunisation to a host, respectively. In particular  
 CC these are useful for ameliorating the symptoms of acquired  
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,

CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
CC compositions are useful for treating or preventing HIV infections or  
CC AIDS. Sequences ABR91449-ABR91624 represent modified HIV protein-encoding  
CC plasmid DNA sequences of the invention.

SO Sequence 6538 BP: 1624 A; 1771 C; 1674 G; 1469 T; 0 other;

#### Alignment Scores:

Pred. No.:	1,21e-24	Length:	6538
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	64.16%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-4 (1-85) x ABR91478 (1-6538)

QY 1 AsnHsThrHrrTrpLeuGluTrpAspArgIuIleAsnTyrrHrSerLeuIleHis 20

DB 3339 AACCAACACCTGGATGGAGTGGGACCGCGAGATCAACAACCTACACCGCTGATCCAC 3398

QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnIuIySaSnGluGlnIuLeuGluLeu 40

DB 3399 AGCTGATCGAGAGAGAGACCAACACGAGAGAAGAAGACGAGAGACTGCTGAGCTG 3458

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52

DB 3459 GACAAGTGGCCAGCTGTGGAACTGTTCAACATC 3494

#### RESULT 13

ID ABR91481 standard; DNA: 6538 BP.

AC ABR91481;

XX 14-AUG-2002 (first entry)

DE Modified HIV protein-encoding plasmid DNA #33.

KM HIV; human immunodeficiency virus; gene: ds; circular; cyclc; anti-HIV;

KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

OS Synthetic.

PN WO200232943-A2.

PD 25-APR-2002.

PF 14-AUG-2001; 2001WO-US25721.

PR 14-AUG-2000; 2000US-225097P.

PR 14-NOV-2000; 2000US-252115P.

PR 28-MAR-2001; 2001US-279257P.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA (CHAD/) CHADRAABARTI B K.

XX Nabel GJ, Huang Y;

XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and

XX its encoded protein, useful as vaccines for genetic or protein

XX immunisation for acquired immunodeficiency syndrome or human

XX immunodeficiency virus infection -

XX Disclosure; Page 449-451; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC these are useful for ameliorating the symptoms of acquired  
CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
CC compositions are useful for treating or preventing HIV infections or  
CC AIDS. Sequences ABR91449-ABR91624 represent modified HIV protein-encoding  
CC plasmid DNA sequences of the invention.

SO Sequence 6538 BP: 1627 A; 1764 C; 1677 G; 1470 T; 0 other;

#### Alignment Scores:

Pred. No.:	1,21e-24	Length:	6538
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	64.16%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-4 (1-85) x ABR91481 (1-6538)

QY 1 AsnHsThrHrrTrpLeuGluTrpAspArgIuIleAsnTyrrHrSerLeuIleHis 20

DB 3339 AACCAACACCTGGATGGAGTGGGACCGCGAGATCAACAACCTACACCGCTGATCCAC 3398

QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnIuIySaSnGluGlnIuLeuGluLeu 40

DB 3399 AGCTGATCGAGAGAGACCAACACGAGAGAAGAAGACGAGAGACTGCTGAGCTG 3458

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52

DB 3459 GACAAGTGGCCAGCTGTGGAACTGTTCAACATC 3494

#### RESULT 14

ID ABR91479 standard; DNA: 6541 BP.

AC ABR91479;

XX 14-AUG-2002 (first entry)

DE Modified HIV protein-encoding plasmid DNA #31.

KM HIV; human immunodeficiency virus; gene: ds; circular; cyclc; anti-HIV;

KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

OS Synthetic.

PN WO200232943-A2.

PD 25-APR-2002.

PF 14-AUG-2001; 2001WO-US25721.

PR 14-AUG-2000; 2000US-225097P.

PR 14-NOV-2000; 2000US-252115P.

PR 28-MAR-2001; 2001US-279257P.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PA (CHAD/) CHADRAABARTI B K.

XX Nabel GJ, Huang Y;

XX WPI; 2002-452382/48.

XX New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and

XX its encoded protein, useful as vaccines for genetic or protein

XX immunisation for acquired immunodeficiency syndrome or human

XX immunodeficiency virus infection -

XX Disclosure; Page 445-447; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV

CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
 CC genetic or protein immunisation to a host, respectively. In particular  
 CC these are useful for ameliorating the symptoms of acquired  
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
 CC compositions are useful for treating or preventing HIV infections or  
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding  
 CC plasmid DNA sequences of the invention.

XX Sequence 6541 BP; 1621 A; 1779 C; 1672 G; 1469 T; 0 other;

#### Alignment Scores:

Pred. No.:	1,21e-24	Length:	6541
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	64.16%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-4 (1-85) x ABK91479 (1-6541)

QY 1 AsnHisThrThrPleuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
 DB 3342 AACCAACACCCCTGATGAGTGGAGCCGCGAGATCAACAACTACACCACTGATCCAC 3401

QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeu 40  
 DB 3402 AGCCTGATCGAGAGAGCCAGAGACAGACAGAGAAAGACAGAGAGCTGCTGGAGCTG 3461

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
 DB 3462 GACAACTGGGCGACCTGTGTGAACTGTTCAACATC 3497

#### RESULT 15

ABK91480  
 ID ABK91480 standard; DNA; 6577 BP.

XX  
 AC ABK91480;

XX  
 DT 14-AUG-2002 (first entry)

XX  
 DE Modified HIV protein-encoding plasmid DNA #32.

XX  
 KW HIV: human immunodeficiency virus; gene; ds: circular; cyclic; anti-HIV;

XX  
 KW Env; Gag; Pol; Nef; genetic immunisation; protein immunisation; AIDS;

XX  
 KW acquired immunodeficiency syndrome; cytotoxic T lymphocyte; CTL.

XX  
 OS Synthetic.

XX  
 PN WO200232943-A2.

XX  
 PD 25-APR-2002.

XX  
 PF 14-AUG-2001; 2001WO-US25721.

XX  
 PR 14-AUG-2000; 2000US-225097P.

XX  
 PR 14-NOV-2000; 2000US-252115P.

XX  
 PR 28-MAR-2001; 2001US-279257P.

XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX  
 PA (CHAD/) CHADRAABARTI B K.

XX  
 PI Nabel GJ, Huang Y;

XX  
 DR WPI; 2002-452382/48.

XX  
 PT New modified human immunodeficiency virus Env, Gag, Pol of Nef DNA and  
 PT its encoded protein, useful as vaccines for genetic or protein  
 PT immunisation for acquired immunodeficiency syndrome or human  
 PT immunodeficiency virus infection  
 XX

PS Disclosure: Page 447-449; 794pp; English.

XX The invention relates to a nucleic acid molecule encoding a modified HIV  
 CC (human immunodeficiency virus) protein. The nucleic acid molecule and its  
 CC encoded protein (modified HIV Env, Gag, Pol or Nef) are useful for  
 CC genetic or protein immunisation to a host, respectively. In particular  
 CC these are useful for ameliorating the symptoms of acquired  
 CC immunodeficiency syndrome (AIDS) or HIV infection and generating an  
 CC antibody or cytotoxic T lymphocyte (CTL) response against native HIV Env,  
 CC Gag, Pol or Nef. Furthermore, the nucleic acids, proteins and subsequent  
 CC compositions are useful for treating or preventing HIV infections or  
 CC AIDS. Sequences ABK91449-ABK91624 represent modified HIV protein-encoding  
 CC plasmid DNA sequences of the invention.

XX Sequence 6577 BP; 1635 A; 1782 C; 1680 G; 1480 T; 0 other;

#### Alignment Scores:

Pred. No.:	1,21e-24	Length:	6577
Score:	290.00	Matches:	51
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	98.08%	Mismatches:	0
Query Match:	64.16%	Indels:	0
DB:	24	Gaps:	0

US-09-877-606-4 (1-85) x ABK91480 (1-6577)

QY 1 AsnHisThrThrPleuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20  
 DB 3378 AACCAACACCCCTGATGAGTGGAGCCGCGAGATCAACAACTACACCACTGATCCAC 3437

QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeuGluLeu 40  
 DB 3438 AGCCTGATCGAGAGAGCCAGAGACAGACAGAGAAAGACAGAGAGCTGCTGGAGCTG 3497

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
 DB 3498 GACAACTGGGCGACCTGTGTGAACTGTTCAACATC 3533

Search completed: March 25, 2003, 13:11:39  
 Job time : 335.626 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:42:33 ; Search time 65.3302 Seconds  
(without alignments)  
399.073 Million cell updates/sec

Title: US-09-877-606-4  
Perfect score: 452  
Sequence: 1 NHTTWLEMDREINNTSLIH.....SKYHENEIARIKKLIGER 85

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 segs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Command line parameters:

-MODE=frame\_p2n.model -DEV=xlh  
-Q/cgn2\_1/USPTO.spool/US09877606/runat\_18032003\_134801\_17715/app.query.fasta\_1.1059  
-DB=Issued\_Patents\_NA -QFMT=fastap -SUFIX=rnl -MINMATCH=0.1 -LOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdl  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09877606.eccn\_1.1.101.@runat\_18032003\_134801\_17715 -NCPU=6 -ICPU=3  
-NO\_XLDPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEDOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

## Database :

- Issued\_Patents\_NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS\_COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	290	64.2	1260	1	US-08-467-933-7 Sequence 7, Appl1
2	290	64.2	1568	1	US-08-589-446-7 Sequence 7, Appl1
3	290	64.2	1568	1	US-08-444-882-7 Sequence 7, Appl1
4	290	64.2	1568	1	US-08-389-459A-7 Sequence 7, Appl1
5	290	64.2	1568	3	US-08-987-867A-7 Sequence 7, Appl1
6	290	64.2	2571	2	US-07-916-098A-1 Sequence 1, Appl1
7	290	64.2	2694	1	US-08-147-890-1 Sequence 1, Appl1
8	290	64.2	3084	1	US-08-147-890-2 Sequence 2, Appl1
9	290	64.2	3084	4	US-08-472-240A-9 Sequence 9, Appl1
10	277	61.3	2644	5	PCT-US95-13335-2 Sequence 2, Appl1
11	277	61.3	1451	1	US-07-924-028A-1 Sequence 1, Appl1
12	277	61.3	1596	3	US-08-388-353-651 Sequence 651, App

13	277	61.3	1596	3	US-08-488-551B-651	Sequence 651, App
14	277	61.3	2730	3	US-08-728-122-1	Sequence 1, Appl1
15	277	61.3	2945	6	5462872-1	Patent No. 5462872
16	277	61.3	3563	3	US-08-463-210-6	Sequence 6, Appl1
17	277	61.3	7399	2	US-08-418-848A-9	Sequence 1, Appl1
18	277	61.3	8932	4	US-09-124-900-1	Sequence 1, Appl1
19	277	61.3	8933	3	US-08-463-210-4	Sequence 4, Appl1
20	277	61.3	8933	4	US-09-620-958A-3	Sequence 3, Appl1
21	277	61.3	8933	4	US-09-620-958A-4	Sequence 4, Appl1
22	277	61.3	8933	4	US-09-620-958A-9	Sequence 9, Appl1
23	277	61.3	9709	2	US-08-188-835-5	Sequence 5, Appl1
24	277	61.3	9709	3	US-08-388-353-1	Sequence 1, Appl1
25	277	61.3	9709	3	US-08-488-551B-1	Sequence 1, Appl1
26	277	61.3	9709	4	US-09-309-572-15	Sequence 15, Appl1
27	277	61.3	12494	4	US-08-935-312-13	Sequence 13, Appl1
28	277	61.3	12494	4	US-08-848-760B-33	Sequence 33, Appl1
29	277	61.3	15581	3	US-08-646-538-35	Sequence 35, Appl1
30	277	61.3	15581	4	US-09-503-222-35	Sequence 35, Appl1
31	275	60.8	2531	4	US-07-956-483-18	Sequence 18, Appl1
32	262	58.0	2184	4	US-09-337-387-4	Sequence 4, Appl1
33	260	57.5	3807	1	US-08-022-835-5	Sequence 5, Appl1
34	260	57.5	3807	1	US-08-388-809-5	Sequence 5, Appl1
35	260	57.5	3807	2	US-08-647-714-5	Sequence 5, Appl1
36	252	55.8	2552	2	US-08-448-603A-27	Sequence 27, Appl1
37	252	55.8	2552	3	US-09-134-075-27	Sequence 27, Appl1
38	252	55.8	2552	2	US-09-492-739-27	Sequence 27, Appl1
39	252	55.8	4527	2	US-08-944-449-8	Sequence 8, Appl1
40	252	55.8	4527	4	US-09-353-362-8	Sequence 8, Appl1
41	244	54.0	342	3	US-08-392-794A-5	Sequence 5, Appl1
42	244	54.0	345	3	US-08-392-794A-20	Sequence 20, Appl1
43	244	54.0	528	3	US-08-392-794A-7	Sequence 7, Appl1
44	244	54.0	531	3	US-08-392-794A-21	Sequence 21, Appl1
45	243	53.8	2571	1	US-08-254-358-3	Sequence 3, Appl1

## ALIGNMENTS

RESULT 1  
US-08-467-933-7  
; Sequence 7, Application US/08467933  
; Patent No. 5736391  
; GENERAL INFORMATION:  
; APPLICANT: Essex, Myron E.  
; APPLICANT: Yu, Xiaofang  
; APPLICANT: Lee, Tun-Hou  
; TITLE OF INVENTION: GP41 MUTANTS AND THEIR USE AS  
; TITLE OF INVENTION: HIV THERAPEUTICS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: WordPerfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467, 933  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/979, 975  
; FILING DATE: NO. 5736391member 23, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29, 066  
; REFERENCE/DOCKET NUMBER: 00379/019001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070

```

; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1260
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
US-08-467-933-7

Alignment Scores:
Pred. No.: 7.17e-28 Length: 1260
Score: 290.00 Matches: 51
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 0
Query Match: 64.16% Indels: 0
DB: 1 Gaps: 0

US-09-877-606-4 (1-85) x US-08-467-933-7 (1-1260)
QY 1 AsnHSThrTTPleuGluTrpAspArgLulLeaAsnAsnYrThrSerleuIleHis 20
Db 464 AATCAGACGACCTGGATGAGATGGACAGAGAAATTAACTTACACAGCTTATATAC 523
QY 21 SerleuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeu 40
Db 524 TCCTTAATTGAAGATCGCAAAACACAGAAAGAAATGAACAAGATTATTGGAATTA 583
QY 41 AsplySTrPAlaSerleuTrpAsnTrpPheAsnIle 52
Db 584 GATAAATGGCAGCTGTGTGGAATTGTTTAACATA 619

RESULT 2
US-08-589-446-7
; Sequence 7, Application US/08589446
; Patent No. 5614413
; GENERAL INFORMATION:
; APPLICANT: Morrow, Casey D.
; TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC
; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/589,446
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/087,009
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Geary III, William C.
; REGISTRATION NUMBER: 31,359
; REFERENCE/DOCKET NUMBER: UAG-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:

```

```

; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1565
US-08-589-446-7

Alignment Scores:
Pred. No.: 9.48e-28 Length: 1568
Score: 290.00 Matches: 51
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 0
Query Match: 64.16% Indels: 0
DB: 1 Gaps: 0

US-09-877-606-4 (1-85) x US-08-589-446-7 (1-1568)
QY 1 AsnHSThrTTPleuGluTrpAspArgLulLeaAsnAsnYrThrSerleuIleHis 20
Db 1264 AATCAGACGACCTGGATGAGATGGACAGAGAAATTAACTTACACAGCTTATATAC 1323
QY 21 SerleuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeu 40
Db 1324 TCCTTAATTGAAGATCGCAAAACACAGAAAGAAATGAACAAGATTATTGGAATTA 1383
QY 41 AsplySTrPAlaSerleuTrpAsnTrpPheAsnIle 52
Db 1384 GATAAATGGCAGCTGTGTGGAATTGTTTAACATA 1419

RESULT 3
US-08-444-882-7
; Sequence 7, Application US/08444882
; Patent No. 5622705
; GENERAL INFORMATION:
; APPLICANT: Morrow, Casey D.
; TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC
; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/444,882
; FILING DATE: 19-MAY-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Geary III, William C.
; REGISTRATION NUMBER: 31,359
; REFERENCE/DOCKET NUMBER: UAG-004
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:

```

NAME/KEY: CDS  
LOCATION: 7..1565  
US-08-444-882-7

Alignment Scores:  
Pred. No.: 9.48e-28 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 1 Gaps: 0

US-09-877-606-4 (1-85) x US-08-444-882-7 (1-1568)

QY 1 AsnHsIhRhrThrPleuGluTrpAspArgGluIleAsnSnyrThrSerLeuIleHis 20  
DB 1264 AATCACACGACCTGGATGAGTGGGACAGAAATTACAACTTACACACTTAATACAC 1323

QY 21 SerLeuIleGluIuSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuLeuLeu 40  
DB 1324 TCCTTAATGAGATCGCAAAACGACAGAAAGAAATGACAGAAATTAATGGAATTA 1383

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
DB 1384 GATTAATGGCAAGTTGTGGAATGTGTTAACATA 1419

RESULT 4  
US-08-389-459A-7  
Sequence 7, Application US/08389459A  
Patent No. 5817512

GENERAL INFORMATION:  
APPLICANT: Morrow, Casey D. and Porter, Donna, C.  
TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT POLIOVIRUS  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
TITLE OF INVENTION: USING SAME  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/389,459A  
FILING DATE: 15-FEB-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/087,009  
FILING DATE: 01-JUL-1993  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Silverl, Jean M.  
REGISTRATION NUMBER: 39,030  
REFERENCE/DOCKET NUMBER: UAG-004CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1568 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7..1565

US-08-389-459A-7

Alignment Scores:  
Pred. No.: 9.48e-28 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 1 Gaps: 0

US-09-877-606-4 (1-85) x US-08-389-459A-7 (1-1568)

QY 1 AsnHsIhRhrThrPleuGluTrpAspArgGluIleAsnSnyrThrSerLeuIleHis 20  
DB 1264 AATCACACGACCTGGATGAGTGGGACAGAAATTACAACTTACACACTTAATACAC 1323

QY 21 SerLeuIleGluIuSerGlnAsnGlnGlnGluLysAsnGlnGlnGluLeuLeuLeu 40  
DB 1324 TCCTTAATGAGATCGCAAAACGACAGAAAGAAATGACAGAAATTAATGGAATTA 1383

QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
DB 1384 GATTAATGGCAAGTTGTGGAATGTGTTAACATA 1419

RESULT 5  
US-08-987-867A-7  
Sequence 7, Application US/08987867A  
Patent No. 6063384

GENERAL INFORMATION:  
APPLICANT: C. Morrow et al.  
TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
TITLE OF INVENTION: USING SAME  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 STATE STREET  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/987,867A  
FILING DATE: 09-DEC-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/087,009  
FILING DATE: 01-JUL-1993  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: UAP-004CPDV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1568 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7..1565  
US-08-987-867A-7

Alignment Scores:

Pred. No.: 9.48e-28 Length: 1568  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 3 Gaps: 0

US-09-877-606-4 (1-85) x US-08-987-867A-7 (1-1568)

OY 1 AsnHstHrThrTPlEgLUtRPaSPaRgGluIleAsnAsnTYrThSeriLeuIleHis 20  
|||||  
DB 1264 AATCAGACGACCTGGATGGAGTGGAGACAGAAATTAACAATTACACAGCTTAATACAC 1323

OY 21 SerLeuIleGluSerGlnAsnGlnGlnGluYsAsnGlnGlnGluLeuLeu 40  
|||||  
DB 1324 TCCTTAATGAAGATCCCAAAACAGACAGAAAGATGAACAGATTAATGGAATTA 1383

OY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
DB 1384 GATAAATGGCAAGTTTGTGGAATGGTTTAACATA 1419

RESULT 6  
US-07-916-098A-1  
Sequence 1, Application US/07916098A  
Patent No. 5871732  
GENERAL INFORMATION:  
APPLICANT: BURKLY, LINDA C.  
APPLICANT: CHISHOLM, PATRICIA L.  
APPLICANT: THOMAS, DAVID W.  
APPLICANT: ROSA, MARGARET D.  
APPLICANT: ROSA, JOSEPH J.  
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN  
TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION  
NUMBER OF SEQUENCES: 61  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.  
STREET: 10 SOUTH WACKER DRIVE  
CITY: CHICAGO  
STATE: ILLINOIS  
COUNTRY: U.S.A.  
ZIP: 60606  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WORD PERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/916,098A  
FILING DATE: July 24, 1992  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/08843  
FILING DATE: No. 5871732ember 27, 1991  
CLASSIFICATION: 424  
APPLICATION NUMBER: 07/618,542  
FILING DATE: No. 5871732ember 27, 1990  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: JOHN J. MC DONNELL  
REGISTRATION NUMBER: 26,949  
REFERENCE/DOCKET NUMBER: 92,310-G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312) 715-1000  
TELEFAX: (312) 715-1234  
TELEX: 910/221-5317  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2571 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
HYPOTHETICAL: NO

ANTI-SENSE: NO  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 1..87  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 88..2568  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1533..1534  
OTHER INFORMATION: /note= "gp120/gp41 cleavage site"

FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..2568  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1  
OTHER INFORMATION: /note= "pre-HIV gp160"

US-07-916-098A-1

Alignment Scores:  
Pred. No.: 1.78e-27 Length: 2571  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 2 Gaps: 0

US-09-877-606-4 (1-85) x US-07-916-098A-1 (1-2571)

OY 1 AsnHstHrThrTPlEgLUtRPaSPaRgGluIleAsnAsnTYrThSeriLeuIleHis 20  
|||||  
DB 1870 AATCAGACGACCTGGATGGAGTGGAGACAGAAATTAACAATTACACAGCTTAATACAC 1929

OY 21 SerLeuIleGluSerGlnAsnGlnGlnGluYsAsnGlnGlnGluLeuLeu 40  
|||||  
DB 1930 TCCTTAATGAAGATCCCAAAACAGACAGAAAGATGAACAGATTAATGGAATTA 1989

OY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
DB 1990 GATAAATGGCAAGTTTGTGGAATGGTTTAACATA 2025

RESULT 7  
US-08-147-890-1  
Sequence 1, Application US/08147890  
Patent No. 5714316  
GENERAL INFORMATION:  
APPLICANT: Weiner, David  
APPLICANT: Williams, William  
APPLICANT: Levy, David N.  
TITLE OF INVENTION: Chimeric Envelope Proteins for Viral  
TITLE OF INVENTION: Targeting  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/147,890  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/718,537  
FILING DATE: 21-JUN-1991  
ATTORNEY/AGENT INFORMATION:



NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST15AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2694 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
US-08-147-890-1

Alignment Scores:  
Pred. No.: 1,89e-27 Length: 2694  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 1 Gaps: 0

US-09-877-606-4 (1-85) x US-08-147-890-1 (1-2694)

Qy 1 AsnHstHrThrTrpLeuGIUTrPAsPArgGluIleAsnAsnTyrrHrSerLeuIleHis 20  
|||||  
Db 1978 AATCAGACGACCTGGATGGAGTGGACAGAGAAATTACACATTAATACAC 2037

Qy 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeu 40  
|||||  
Db 2038 TCCTTATTGAGAAATCGCAAAACGCAAGAAAGATGACAGATATTATTGGAATTA 2097

Qy 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 2098 GATAAATGGCGCAAGTTGTGGAATGCTTAACATA 2133

RESULT 8  
US-08-147-890-2  
Sequence 2, Application US/08147890  
Patent No. 5714316  
GENERAL INFORMATION:  
APPLICANT: Weiner, David  
APPLICANT: Williams, William  
APPLICANT: Levy, David N.  
TITLE OF INVENTION: Chimeric Envelope Proteins for Viral  
TIME OF INVENTION: Targeting  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/147,890  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/718,537  
FILING DATE: 21-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST15AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200

TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3084 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
US-08-147-890-2

Alignment Scores:  
Pred. No.: 2,24e-27 Length: 3084  
Score: 290.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 98.08% Mismatches: 0  
Query Match: 64.16% Indels: 0  
DB: 1 Gaps: 0

US-09-877-606-4 (1-85) x US-08-147-890-2 (1-3084)

Qy 1 AsnHstHrThrTrpLeuGIUTrPAsPArgGluIleAsnAsnTyrrHrSerLeuIleHis 20  
|||||  
Db 2368 AATCAGACGACCTGGATGGAGTGGACAGAGAAATTACACATTAATACAC 2427

Qy 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGluGlnGluLeuLeu 40  
|||||  
Db 2428 TCCTTATTGAGAAATCGCAAAACGCAAGAAAGATGACAGATATTATTGGAATTA 2487

Qy 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52  
|||||  
Db 2488 GATAAATGGCGCAAGTTGTGGAATGCTTAACATA 2523

RESULT 9  
US-08-472-240A-9  
Sequence 9, Application US/08472240A  
Patent No. 6284248  
GENERAL INFORMATION:  
APPLICANT: KIENY, Marie-Paule  
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
TITLE OF INVENTION: gp160 VARIANT  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/472,240A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/956,483  
FILING DATE: 31-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 017753-055  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2644 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant

```

; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 12..2600
US-08-472-240A-9

Alignment Scores:
Pred. No.: 6, 11e-26      Length: 2644
Score: 278.00      Matches: 51
Percent Similarity: 78.67%      Conservative: 8
Best Local Similarity: 68.00%      Mismatches: 4
Query Match: 61.50%      Indels: 12
DB: 4      Gaps: 1

US-09-877-606-4 (1-85) x US-08-472-240A-9 (1-2644)

QY 1 AsnHsThrThrTrrpleuGlutrrpAspArgGlulIeasnaSnYrrThrSerleuIleHis 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db: 1896 AATTAATGACCTGCTGATGGAGTGGGACAGAGAAATTAAACATTACACAACTTAATACAT 1955

QY 21 SerleuIleGlulIeuserIleSerIleAsnGlnGlnIulysAsnGlnGlnIleuLeu 40
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db: 1956 TCGTTAATGAGAAATTCGCAAAACACAGAAAGAAATGAACAAGATTATTGGAATTA 2015

QY 41 AspIlyStrAlaSerleuTrpAsnTrpPheAsnIleLysIleLysGlnIleGluAspLys 60
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db: 2016 GATTAATGGCGCAAGTTTGTCGAATGTGTTAAC----- 2048

QY 61 IleGluGluIleuserIleTyrrHisIleGluAsnGluIle 75
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db: 2049 -----ATAACAATTTGCTGTGTTATATAAAAAATAGAGTT 2084

RESULT 10
PCT-US95-13335-2
; Sequence-2, Application PC/TUS9513335
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A RECOMBINANT PROTEIN USEFUL IN THE
; TITLE OF INVENTION: DETECTION OF HIV, DNA SEQUENCES ENCODING THE PROTEIN, AND
; TITLE OF INVENTION: IMMUNOASSAYS USING THE PROTEIN.
; NUMBER OF SEQUENCES: 2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13335
; FILING DATE:
; CLASSIFICATION:
; INFORMATION FOR SEO ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 699 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
PCT-US95-13335-2

Alignment Scores:
Pred. No.: 1, 5e-26      Length: 699
Score: 277.00      Matches: 49
Percent Similarity: 98.08%      Conservative: 2
Best Local Similarity: 94.23%      Mismatches: 1
Query Match: 61.28%      Indels: 0
DB: 5      Gaps: 0

US-09-877-606-4 (1-85) x PCT-US95-13335-2 (1-699)

QY 1 AsnHsThrThrTrrpleuGlutrrpAspArgGlulIeasnaSnYrrThrSerleuIleHis 20
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db: 388 AATTAATGACCTGCTGATGGAGTGGGACAGAGAAATTAAACATTACACAACTTAATACAT 447

```

```

0Y      21 SerLeuIIeGIuGIuSerGIuAsnGIuInGIuInGIuLyAsnGIuInGIuLeuGIuLeu 40
Db      448 TCCTTAATTGAAGAATCGCAAAACACGACAGAAAAGATGACAAAGATTATTGGAAATTA 507
0Y      41 AspLySTrPalaseLerLeuTrpAsnTrpPheAsnIle 52
Db      508 GATTAATGGCGCAAGTTTGTGGAATTGGTTTAAACATA 543

RESULT 11
US-07-924-028A-1
; Sequence 1, Application US/07924028A
; Patent No. 5470573
; GENERAL INFORMATION:
; APPLICANT: Lubitz Werner, Szostak, Michael P.
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/924,028A
; FILING DATE: 30-SEP-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP91/00308
; FILING DATE: 02-FEB-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 40 05 874
; FILING DATE: 24-FEB-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5470573man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: HUBR 1027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1451 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-07-924-028A-1

Alignment Scores:
Pred. No.: 3.81e-26 Length: 1451
Score: 277.00 Matches: 49
Percent Similarity: 98.08% Conservative: 2
Best local Similarity: 94.23% Mismatches: 1
Query Match: 61.28% Indels: 0
DB: 1 Gaps: 0

US-09-877-606-4 (1-85) x US-07-924-028A-1 (1-1451)
0Y      1 AsnHISThrTrpLeuGIuTrpAspArgSLuIleAsnAsnTrpThrSerLeuIleHis 20
Db      496 AATAACATGACCTCGATGCGAGTGGGACAGACAATAATTACATTTACACAGCTTAATACAC 555
0Y      21 SerLeuIIeGIuGIuSerGIuAsnGIuInGIuInGIuLyAsnGIuInGIuLeuGIuLeu 40
Db      556 TCCTTAATTGAAGAATCGCAAAACACGACAGAAAAGATGACAAAGATTATTGGAAATTA 615
0Y      41 AspLySTrPalaseLerLeuTrpAsnTrpPheAsnIle 52

```

```

Db      616   GATAAATGGCAAGTTTGTGGAATTGCTTTAACATA    651
|||||
RESULT 12
US-08-388-353-651
; Sequence 651, Application US/08388353
; Patent No. 6010895
GENERAL INFORMATION:
APPLICANT: Deacon, Nicholas J.
APPLICANT: Leamont, Jennifer C.
APPLICANT: McPhee, Dale A.
APPLICANT: Crowe, Suzanne
APPLICANT: Cooper, David
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
NUMBER OF SEQUENCES: 800
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: United States
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/388,353
FILING DATE: 14-FEB-1995
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S.
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 9606
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 742-4343
TELEFAX: (516) 742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 651:
SEQUENCE CHARACTERISTICS:
LENGTH: 1596 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-388-353-651

Alignment Scores:
Pred. NO.:          4.3e-26           Length:       1596
Score:              277.00            Matches:       49
Percent Similarity: 98.08%            Conservative: 2
Best local Similarity: 94.23%          Mismatches:   1
Query Match:        61.28%             Indels:        0
DB:                  3                gaps:         0

US-09-877-606-4 (1-85) x US-08-388-353-651 (1-1596)
QY      1     ASNHSHTRHRTPLRGLUTTPASPARGLUTLNASNNTYRRHRserleullehis 20
|||
Db      13     AAATACATGACCCTGGATGGATGGATGGACAGAGAATTAACAATTACAAACCTTAATTCAC 72
|||
QY      21     Serleullelglugluserglasnglnglnlugysasngluglnglnuleleugluieu 40
|||
Db      73     TCCCTAANTGAAGAATGCCAANAACACAGCAAGAAAAGATGACACAGATTAATTGGAAATTA 132
|||
QY      41     AsplystirpalaserLeuttpAsnttPhehsnile 52
|||
Db      133    GATAAATGGCGCAAGTTGTGTGAATTCGTTTAACATA 168
|||
RESULT 13
US-08-488-551B-651
; Sequence 651, Application US/08488551B

```

```

      : Patent No.6015661
      : General Information:
      : APPLICANT: Nicholas J. Deacon
      : APPLICANT: Dale A. McPhee
      : APPLICANT: David Cooper
      : TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1
      : NUMBER OF SEQUENCES: 841
      : CORRESPONDENCE ADDRESS:
      : ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER
      : STREET: 400 GARDEN CITY PLAZA
      : CITY: GARDEN CITY
      : STATE: NEW YORK
      : COUNTRY: U.S.A.
      : ZIP: 11530-0299
      : COMPUTER READABLE FORM:
      : MEDIUM TYPE: Floppy disk
      : COMPUTER: IBM PC compatible
      : OPERATING SYSTEM: PC-DOS/MS-DOS
      : SOFTWARE: Patentln Release #1.0, Version #1.25
      : CURRENT APPLICATION DATA:
      : APPLICATION NUMBER: US/08/488,551B
      : FILING DATE: 07-JUN-1995
      : PRIOR APPLICATION DATA:
      : APPLICATION NUMBER: PM3864 (AU)
      : FILING DATE: 14-FEB-1994
      : APPLICATION NUMBER: PM4002 (AU)
      : FILING DATE: 21-FEB-1994
      : APPLICATION NUMBER: PN0284 (AU)
      : FILING DATE: 23-DEC-1994
      : APPLICATION NUMBER: US 08/388,353
      : FILING DATE: 14-FEB-1995
      : APPLICATION NUMBER: PN3021/95
      : FILING DATE: 17-MAY-1995
      : ATTORNEY/AGENT INFORMATION:
      : NAME: FRANK S. DIGILIO
      : REFERENCE/DOCKET NUMBER: 9606Z
      : TELECOMMUNICATION INFORMATION:
      : TELEPHONE: (516) 742-4343
      : TELEFAX: (516) 742-4366
      : INFORMATION FOR SEQ ID NO: 651:
      : SEQUENCE CHARACTERISTICS:
      : LENGTH: 1596 base pairs
      : TYPE: nucleic acid
      : STRANDEDNESS: single
      : TOPOLOGY: linear
      : MOLECULE TYPE: DNA
      :
      : US-08-488-551B-651
      :
      Alignment Scores:
      Pred. No.:
      Score: 4.3e-26 Length: 1596
      Percent Similarity: 277.00 Matches: 49
      Best local Similarity: 98.08% Conservative: 2
      Query Match: 61.28% Mismatches: 1
      DB: 3 Indels: 0
      Gaps: 0
      :
      US-09-877-606-4 (1-85) x US-08-488-551B-651 (1-1596)
      QY 1 ASNHSTHRTTTPLEUGLUTIPASARGLUITLKSANSNYRTHRSerLeuIleHis 20
      Db 13 AAATACATGAGCTGGATGGATGGACGAGAATAATTACAATTTCACAGACTTAATCAC 72
      QY 21 SerLeuIleGluInuSerGlnAsnGluInGluLyAsnGluInGluLeuLeuGluLeu 40
      Db 73 TCCTTAAATGAAGATGCACAACACAGCAAGAAAAGATGACACAGAAATTATTGGAATT 132
      QY 41 AsplystrPalaserLeuTPasnttPpheksnIle 52
      Db 133 GATAAATGGCGCAAGCTTGTCGAATTGGTTAACATA 168
      RESULT 14
      : Sequence 1, Application US/08728122

```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 25, 2003, 15:49:42 ; Search time 120.172 Seconds  
(without alignments)  
548.507 Million cell updates/sec

Title: US-09-877-606-4

Perfect score: 452

Sequence: 1 NHTWLEMDREINNTSLIH.....SKYIENEIARIKILIGER 85

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 538826 seqs, 387737923 residues

Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=ylh  
-Q/cgcn2\_1/USFT0.spool/US09877606/runat\_18032003\_134802\_17783/app.query.fasta\_1.1059  
-DB=Published.Applications.NA -QEXT=fastap -SUFF=trpb -MINMATCH=0.1  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100  
-THR.MN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09877606 -QCGN\_1.1.115 -runat\_18032003\_134802\_17783  
-KPU=6 -ICPU=3 -NO.XLPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG  
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA:\*

1: /cgcn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgcn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgcn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgcn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgcn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgcn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgcn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgcn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgcn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgcn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgcn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgcn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgcn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgcn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	64.2	1568	10	US-09-756-551A-7
2	290	64.2	9719	10	US-09-737-190A-1
3	277	61.3	2280	9	US-10-003-035-16
4	277	61.3	2562	9	US-10-003-035-32

5	277	61.3	2583	9	US-10-003-035-24	Sequence 24, Appl
6	277	61.3	2746	9	US-10-003-035-20	Sequence 20, Appl
7	277	61.3	2747	9	US-10-003-035-23	Sequence 23, Appl
8	277	61.3	2941	9	US-10-003-035-19	Sequence 19, Appl
9	277	61.3	2950	9	US-10-003-035-22	Sequence 22, Appl
10	277	61.3	3157	9	US-10-003-035-14	Sequence 14, Appl
11	277	61.3	3417	9	US-10-003-035-21	Sequence 21, Appl
12	277	61.3	3839	9	US-10-003-035-52	Sequence 52, Appl
13	277	61.3	4040	9	US-10-003-035-54	Sequence 54, Appl
14	277	61.3	8933	10	US-09-943-286-3	Sequence 3, Appl
15	277	61.3	8933	10	US-09-943-286-4	Sequence 4, Appl
16	277	61.3	8933	10	US-09-943-286-9	Sequence 9, Appl
17	277	61.3	9918	10	US-09-798-675-5	Sequence 5, Appl
18	252	55.8	2298	10	US-09-476-242-22	Sequence 22, Appl
19	252	55.8	2298	10	US-09-476-242-23	Sequence 23, Appl
20	252	55.8	2298	10	US-09-476-242-24	Sequence 24, Appl
21	252	55.8	2310	10	US-09-476-242-3	Sequence 3, Appl
22	252	55.8	2310	10	US-09-476-242-21	Sequence 21, Appl
23	252	55.8	2316	10	US-09-476-242-4	Sequence 4, Appl
24	252	55.8	2316	10	US-09-476-242-8	Sequence 8, Appl
25	252	55.8	2332	10	US-09-476-242-5	Sequence 5, Appl
26	252	55.8	2332	10	US-09-476-242-18	Sequence 18, Appl
27	252	55.8	2332	10	US-09-476-242-19	Sequence 19, Appl
28	252	55.8	2332	10	US-09-476-242-20	Sequence 20, Appl
29	252	55.8	2338	10	US-09-476-242-6	Sequence 6, Appl
30	252	55.8	2334	10	US-09-476-242-7	Sequence 7, Appl
31	252	55.8	2352	10	US-09-476-242-26	Sequence 26, Appl
32	252	55.8	2358	10	US-09-476-242-25	Sequence 25, Appl
33	252	55.8	2358	10	US-09-476-242-16	Sequence 16, Appl
34	252	55.8	2517	10	US-09-476-242-17	Sequence 17, Appl
35	252	55.8	2523	10	US-09-476-242-15	Sequence 15, Appl
36	252	55.8	2529	10	US-09-476-242-14	Sequence 14, Appl
37	252	55.8	2535	10	US-09-476-242-13	Sequence 13, Appl
38	252	55.8	2541	10	US-09-476-242-9	Sequence 9, Appl
39	252	55.8	2541	10	US-09-476-242-10	Sequence 10, Appl
40	252	55.8	2541	10	US-09-476-242-11	Sequence 11, Appl
41	252	55.8	2541	10	US-09-476-242-12	Sequence 12, Appl
42	251	55.5	1860	10	US-09-891-609-3	Sequence 3, Appl
43	251	55.5	1941	10	US-09-891-609-1	Sequence 1, Appl
44	243	53.8	2571	9	US-10-077-294-3	Sequence 3, Appl
45	243	53.8	2571	9	US-10-163-886-3	Sequence 3, Appl

#### ALIGNMENTS

RESULT 1  
US-09-756-551A-7  
; Sequence 7, Application US/09756551A  
; Patent No. US20020051768A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Morrow et al.  
; TITLE OF INVENTION: ENCAPSATED RECOMBINANT VIRAL  
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 28 STATE STREET  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/756,551A  
; FILING DATE: 08-JAN-2001  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/376,184

```

: FILING DATE: 17-AUG-1999
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/987,867
: FILING DATE: 09-DEC-1997
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/389,459
: FILING DATE: 15-FEB-1995
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/087,009
: FILING DATE: 01-JUL-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Lauro, Peter C.
: REGISTRATION NUMBER: 32,360
: REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (617) 227-7400
: TELEFAX: (617) 742-4214
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1568 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 7..1565
US-09-756-551A-7

Alignment Scores:
Pred. No.: 3.78e-27 Length: 1568
Score: 290.00 Matches: 51
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 0
Query Match: 64.16% Indels: 0
DB: 10 Gaps: 0

US-09-877-606-4 (1-85) x US-09-756-551A-7 (1-1568)
QY 1 AsnHsthrTrpLeuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20
: |||||
Db 1264 AATCAGACGACCTGGATGGATGGACAGACAATTAACATTACACAGCTTAATACAC 1323
QY 21 SerLeuIleGluSerGlnAsnGlnGlnIuLysAsnGlnGluLeuLeu 40
: |||||
Db 1324 TCCTTAATTGAAGATGCCAAACACAGCAAGAAAGATGAACAAGATTATTGGAATTA 1383
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
: |||||
Db 1384 GATAAATGGCGCAAGTTGTGGAATTGTTAAACATA 1419

RESULT 2
US-09-737-190A-1
: Sequence 1, Application US/09737190A
: Patent No. US20020102545A1
: GENERAL INFORMATION:
: APPLICANT: Shibusu, Tetsuo
: TITLE OF INVENTION: A Method for Changing a Target Array, a Method for Analyzing
: TITLE OF INVENTION: a Structure, and an Apparatus, a Storage Medium and a
: FILE REFERENCE: JP919902700S1 (14043)
: CURRENT APPLICATION NUMBER: US/09/737,190A
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 1
: LENGTH: 9719
: TYPE: DNA
: ORGANISM: Human Immunodeficiency virus type 1
US-09-737-190A-1
```

```

Alignment Scores:
Pred. No.: 3.98e-26 Length: 9719
Score: 290.00 Matches: 51
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 98.08% Mismatches: 0
Query Match: 64.16% Indels: 0
DB: 10 Gaps: 0

US-09-877-606-4 (1-85) x US-09-737-190A-1 (1-9719)
QY 1 AsnHsthrTrpLeuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20
: |||||
Db 8094 AATCAGACGACCTGGATGGATGGACAGACAATTAACATTACACAGCTTAATACAC 8153
QY 21 SerLeuIleGluSerGlnAsnGlnGlnIuLysAsnGlnGluLeuLeu 40
: |||||
Db 8154 TCCTTAATTGAAGATGCCAAACACAGCAAGAAAGATGAACAAGATTATTGGAATTA 8213
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
: |||||
Db 8214 GATAAATGGCGCAAGTTGTGGAATTGTTAAACATA 8249

RESULT 3
US-10-003-035-16
: Sequence 16, Application US/10003035
: Patent No. US2002015127A1
: GENERAL INFORMATION:
: APPLICANT: Wang, Danher
: TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
: FILE REFERENCE: 22488-712
: CURRENT APPLICATION NUMBER: US/10/003,035
: PRIOR FILING DATE: 2001-11-01
: PRIOR APPLICATION NUMBER: 09/585,599
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US01/18238
: PRIOR FILING DATE: 2001-06-04
: NUMBER OF SEQ ID NOS: 75
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 16
: LENGTH: 2280
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: Modified Env
US-10-003-035-16

Alignment Scores:
Pred. No.: 2.58e-25 Length: 2280
Score: 277.00 Matches: 49
Percent Similarity: 98.08% Conservative: 2
Best Local Similarity: 94.23% Mismatches: 1
Query Match: 61.28% Indels: 0
DB: 9 Gaps: 0

US-09-877-606-4 (1-85) x US-10-003-035-16 (1-2280)
QY 1 AsnHsthrTrpLeuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20
: |||||
Db 1876 AATAACATGACCTGGATGGATGGACAGACAAGAAATTAAACATTACACAACTTAATACAC 1935
QY 21 SerLeuIleGluSerGlnAsnGlnGlnIuLysAsnGlnGluLeuLeu 40
: |||||
Db 1936 TCCTTAATTGAAGATGCCAAACACAGCAAGAAAGATGAACAAGATTATTGGAATTA 1995
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
: |||||
Db 1996 GATAAATGGCGCAAGTTGTGGAATTGTTAAACATA 2031

RESULT 4
US-10-003-035-32
: Sequence 32, Application US/10003035
: Patent No. US2002015127A1
: GENERAL INFORMATION:
```

Alignment Scores:	
Pred. No.:	3, 04e-25
Score:	277.00
Percent Similarity:	98.08%
Best Local Similarity:	94.22%
Query Match:	61.28%
DB:	9
	Gaps: 0
	length: 2583
	Matches: 49
	Conservative: 2
	Mismatches: 1
	Indels: 0
	Gaps: 0

PRIOR APPLICATION NUMBER: PCT/US01/18238

```

: PRIOR FILING DATE: 2001-06-04
: NUMBER OF SEQ ID NOS: 75
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 23
: LENGTH: 2747
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: Modified Env/Tat from strain BH10
US-10-003-035-23

Alignment Scores:
Pred. No.:
Score: 3.29e-25 Length: 2747
Percent Similarity: 277.00 Matches: 49
Best Local Similarity: 98.088 Conservative: 2
Query Match: 94.238 Mismatches: 1
DB: 61.284 Indels: 0
Gaps: 0

US-09-877-606-4 (1-85) x US-10-003-035-23 (1-2747)

OY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20
Db 2343 AATAACTGACCTGGATGGAGTGGGACAGAGAATTAACAATTACACAAGCTTATACAC 2402
OY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluIleGluIleGluIleGluIle 40
Db 2403 TCCTTAATTGAAGATGCGCAAAACCGACAGAAAGAAAGAAATGACAGAAATATTGGAAATTA 2462
OY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
Db 2463 GATAAATGGCAAGTTTGCTGGAATGGTTTAACATA 2498

RESULT 8
US-10-003-035-19
: Sequence 19, Application US/10003035
: Patent No. US20020155127A1
: GENERAL INFORMATION:
: APPLICATION: Wang, Danher
: TITLE OR INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
: FILE REFERENCE: 22488-712
: CURRENT APPLICATION NUMBER: US/10/003,035
: PRIOR FILING DATE: 2001-11-01
: PRIOR APPLICATION NUMBER: 09/585,599
: PRIOR FILING DATE: 2000-06-02
: PRIOR APPLICATION NUMBER: PCT/US01/18238
: PRIOR FILING DATE: 2001-06-04
: NUMBER OF SEQ ID NOS: 75
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 19
: LENGTH: 2941
: TYPE: DNA
: ORGANISM: Artificial sequence
: FEATURE:
: OTHER INFORMATION: Modified Env from HIV strain pNL4-3
US-10-003-035-19

Alignment Scores:
Pred. No.:
Score: 3.59e-25 Length: 2941
Percent Similarity: 277.00 Matches: 49
Best Local Similarity: 98.088 Conservative: 2
Query Match: 94.238 Mismatches: 1
DB: 61.284 Indels: 0
Gaps: 0

US-09-877-606-4 (1-85) x US-10-003-035-19 (1-2941)

OY 1 AsnHisThrTrpLeuGluTrpAspArgGluIleAsnAsnTyrThrSerLeuIleHis 20
Db 2336 AATAACTGACCTGGATGGAGTGGGACAGAGAATTAACAATTACACAAGCTTATACAC 2395
OY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluIleGluIleGluIleGluIle 40

```

```
Db      2396   TCCTTAATGCAAGATCGCAAACACGACAAGAAGAAGAAATGAACAGAAATTATTTGAATTA    2455
QY       41 Asplystirpalaserleutripasnttppheasnille 52
          |||||||
Db      2456   GATAAATGGCGCAAGTTGTGTGGAATTGCTTAAACATA    2491
RESULT 9
US-10-003-035-22
; Sequence 22, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Mang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
FILE REFERENCE: 22488-712
CURRENT APPLICATION NUMBER: US/10/003,035
CURRENT FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: 09/585,599
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US01/18238
PRIOR FILING DATE: 2001-06-04
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 22
LENGTH: 2950
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Modified Env/Nef from strain BH10
US-10-003-035-22
Alignment Scores:
Pred. No.:           3.6e-25             Length:         2950
Score:              277.00               Matches:         49
Percent Similarity: 98.08%               Conservative:     2
Best Local Similarity: 94.23%             Mismatches:      1
Query Match:        61.28%                Indels:          0
DB:                 9                     Gaps:            0
US-09-877-606-4 (1-85) x US-10-003-035-22 (1-2950)
QY       1 AsnHisThrThrpLeugluTrpPaspArgJulIeAsnAsnTyThrSerLeuIleHis 20
          |||||: |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      1876   AATAACATGACTGCATGCATGCAGTGCGACACAGAAATTAACAAATTACACAAGCTTATATCAC 1935
QY       21 SerLeuIleGlusJerGinsngIngIngIuLySasngIngIngIuleuLeuGluLeu 40
          |||||||
Db      1936   TCCTTATTGAAAGATTCGCAAAACCGACGAAGAAGAAAGATGAACAGAAATTATGGAATTA 1995
QY       41 Asplystirpalaserleutripasnttppheasnille 52
          |||||||
Db      1996   GATAAATGGCGCAAGTTGTGTGGAATTGCTTAAACATA    2031
RESULT 10
US-10-003-035-14
; Sequence 14, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Mang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
FILE REFERENCE: 22488-712
CURRENT APPLICATION NUMBER: US/10/003,035
CURRENT FILING DATE: 2001-11-01
PRIOR APPLICATION NUMBER: 09/585,599
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US01/18238
PRIOR FILING DATE: 2001-06-04
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 3157
TYPE: DNA
ORGANISM: Artificial sequence
FEATURE:
```





Query Match:	61.28%	Indels:	0
DB:	9	Gaps:	0
US-09-877-606-4 (1-85) x US-10-003-035-54 (1-4040)			
QY	1	AsnHisThrTrpLeuGluTrpAspArgLulLeaAsnAsnTyrThrSerLeuIleHis	20
DB	3435	AATACATGACCTGCATGCGATGGCGACAGAAATTTACCAATTACACAGCTTAATACAC	3499
QY	21	SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGlnLeuLeuGluLeu	40
DB	3495	TCCCTAATTGCAAGATGCGCAAAACCGACAGAAAGAAAGAACAGAAATTTGGAATTA	3554
QY	41	AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle	52
DB	3555	GATAAATGGCGAAGTTGTGTGAATTTGTTTAACATA	3590
RESULT 14			
US-09-943-286-3			
; Sequence 3, Application US/09943286			
; Patent No. US2002010668A1			
; GENERAL INFORMATION:			
; APPLICANT: Nunomura, Kiyolada			
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD			
; FILE REFERENCE: GP104-02 JP			
; CURRENT APPLICATION NUMBER: US/09/943,286			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: FastSeq for Windows Version 3.0			
; SEQ ID NO 3			
LENGTH: 8933			
TYPE: RNA			
ORGANISM: Human Immunodeficiency Virus			
FEATURE:			
NAME/KEY: source			
LOCATION: (1)...(8933)			
OTHER INFORMATION: Sequence of transcripts produced from the BH10			
US-09-943-286-3			
Alignment Scores:			
Pred:	1.51e-24	Length:	8933
Score:	227.00	Matches:	49
Percent Similarity:	98.08%	Conservative:	2
Best local Similarity:	94.23%	Mismatches:	1
Query Match:	61.28%	Indels:	0
DB:	10	Gaps:	0
US-09-877-606-4 (1-85) x US-09-943-286-3 (1-8933)			
QY	1	AsnHisThrTrpLeuGluTrpAspArgLulLeaAsnAsnTyrThrSerLeuIleHis	20
DB	7450	AATACATGACCTGCATGCGATGGCGACAGAAATTTACCAATTACACAGCTTAATACAC	7509
QY	21	SerLeuIleGluGluSerGlnAsnGlnGlnGluLysAsnGlnGlnLeuLeuGluLeu	40
DB	7510	UCCCUAUAUUGAAGAACCGCAAAACCGACAGAAAGAAAGAACAGCAAGAAUUAUUGGAUUA	7569
QY	41	AspLysTrpAlaSerLeuTrpAsnTrpPheAsnIle	52
DB	7570	GAUAAAUGGGCAAGUUGUGGAAUUGGUUUAACATA	7605
RESULT 15			
US-09-943-286-4			
; Sequence 4, Application US/09943286			
; Patent No. US2002010668A1			
; GENERAL INFORMATION:			
; APPLICANT: Nunomura, Kiyolada			
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD			
; FILE REFERENCE: GP104-02 JP			
; CURRENT APPLICATION NUMBER: US/09/943,286			
; FILING DATE: 2001-08-30			
NUMBER OF SEQ ID NOS: 9			

```

: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 4
: LENGTH: 8933
: TYPE: RNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Sequence of the IAC-Asrc pseudo target
: NAME/KEY: mutation
: LOCATION: (4135)...(4155)
: OTHER INFORMATION: Mutated positions: 4135, 4140-1, 4145, 4150,
: OTHER INFORMATION: 4152-3, 4155
: US-09-943-286-4

Alignment Scores:
Pred. No.: 1 51c-24 8933
Score: 277.00
Percent Similarity: 98.08%
Best Local Similarity: 94.23%
Query Match: 61.28%
DB: 10 Gaps: 0

US-09-877-606-4 (1-85) x US-09-943-286-4 (1-8933)

QY 1 AsnHisThrThrpLeuGluTrpAspArgIuIleAsnAsnTrpThrSerLeuIleHis 20
||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 7450 AAUAACUGAGACGUGAGUGGAGUGGACAGACAAAUUAACAAUUAACACAGCUAAUACAC 7509

QY 21 SerLeuIleGluGluSerGlnAsnGlnGlnGluIuLyAsnGluGlnGluLeuGluLeu 40
||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 7510 UCCUUAUUGGAAUAUCCGCAAAACCGACAGCAAGAAAGAAUGAGACAGAAUUAUUGGAUUA 7569

QY 41 AspLyfTrpAlaSerLeuTrpAsnTrpPheAsnIle 52
||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||: |||||:
Db 7570 GAUAAUUGGCAAGUUGUGGAUUGGUUUAACAAUA 7605

```

```
Search completed: March 25, 2003, 20:14:57
Job time : 124.172 secs
```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:36 ; Search time 27.2167 Seconds  
(without alignments)  
300.235 Million cell updates/sec

Title: US-09-877-606-4

Perfect score: 452

Sequence: 1 NHTTWLEMDREINNYTSLIH.....SKYHENEIARIKKLIGER 85

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

1: PIR\_73:\*  
2: pir1:\*  
3: pir2:\*  
4: pir3:\*  
5: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	277	61.3	851	2 S33985	env polyprotein -
2	277	61.3	854	2 S13288	env protein - huma
3	277	61.3	856	1 VCLJH3	env polyprotein pr
4	277	61.3	861	1 VCLJLV	env polyprotein pr
5	268	59.3	856	1 VCLJLV	env polyprotein pr
6	261	57.7	358	2 S21998	envelope protein g
7	254	56.2	856	1 VCLJ3W	env polyprotein pr
8	253	56.0	861	1 VCLJSC	env polyprotein pr
9	231	55.5	443	2 C41621	env polyprotein p
10	247	54.6	847	2 T09448	envelope glycoprot
11	247	54.6	847	2 S13289	env protein - huma
12	246	54.4	852	2 T12016	envelope glycoprot
13	245	54.2	357	2 S21996	envelope protein g
14	244	54.0	358	2 S22002	envelope protein g
15	244	54.0	358	2 S22000	envelope protein g
16	244	54.0	358	2 S70417	envelope protein g
17	243	53.8	859	1 VCLJMN	env polyprotein pr
18	242	53.5	357	2 S22006	envelope protein g
19	241	53.3	357	2 S21994	envelope protein g
20	241	53.3	853	2 S54384	envelope protein g
21	240	53.3	855	1 VCLJZR	env polyprotein pr
22	240	53.1	357	2 S21992	envelope protein g
23	239	52.9	852	1 VCLJBR	env polyprotein -
24	237	52.4	357	2 S22004	envelope protein g
25	237	52.4	445	2 A41621	env polyprotein M
26	236	52.2	729	1 VCLJAX	env polyprotein pr
27	236	52.2	855	1 VCLJAZ	env polyprotein pr
28	236	52.2	861	1 VCLJKB	env polyprotein pr
29	236	52.2	868	1 VCLJH4	env polyprotein -

30	234	51.8	846	1 VCLJND	env polyprotein pr
31	233	51.5	843	1 H44001	env polyprotein pr
32	230	50.9	859	2 T01672	envelope polyprote
33	228	50.4	136	2 JU0266	envelope polyprote
34	228	50.4	454	2 B41621	env polyprotein D
35	223	49.3	136	2 JT0954	envelope polyprote
36	215	47.6	856	1 A44963	env polyprotein pr
37	213	47.1	357	2 S21990	envelope protein g
38	198	43.8	854	1 VCLJST	env polyprotein pr
39	197	43.6	877	2 A59197	envelope protein p
40	181	40.0	863	2 A53034	gag polyprotein -
41	137.5	30.4	881	1 VCLJG3	env polyprotein -
42	136	30.1	281	1 R6BYA2	amino acid biosynt
43	133.5	29.5	151	2 S30448	env protein - huma
44	133.5	29.5	151	2 S30452	env protein - huma
45	133.5	29.5	881	2 S03068	env protein - huma

#### ALIGNMENTS

##### RESULT 1

S33985 env polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999

C:Accession: S33985

R:Carlini, F.  
submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33985

A>Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-851 <CAR>

A:Cross-references: EMBL:211530; NID:960192; PIDN:CAA77628.1; PID:960199

C:Superfamily: type E retrovirus env polyprotein

Query Match 61.3% Score 277; DB 2; Length 851;

Best Local Similarity 94.2% Pred. No. 4.4e-16; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NHTTWLEMDREINNYTSLIHSLIEESONOEKNEOELLELDKMASLNMWNI 52  
Db 619 NNTTWMEDEINNYTSLIHSLIEESONOEKNEOELLELDKMASLNMWNI 670

##### RESULT 2

S13288 env protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997

C:Accession: S13288

R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, Nature 348, 69-73, 1990

A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp1

A:Reference number: S13288; MID:91043044; PMID:2172833

A:Accession: S13288

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-854 <OBR>

C:Superfamily: type E retrovirus env polyprotein

Query Match 61.3% Score 277; DB 2; Length 854;

Best Local Similarity 94.2% Pred. No. 4.4e-16; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NHTTWLEMDREINNYTSLIHSLIEESONOEKNEOELLELDKMASLNMWNI 52  
Db 622 NNTTWMEDEINNYTSLIHSLIEESONOEKNEOELLELDKMASLNMWNI 673

RESULT 3  
VCLJH3

env polyprotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03973  
R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Doranberger, J.A.; Papas, T.S.; Ghrayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.  
Nature 313, 277-284, 1985  
A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.  
A:Reference number: A93353; MUID:85111123; PMID:2578615  
A:Accession: A03973  
A:Molecule type: DNA  
A:Residues: 1-856 <RAT>  
A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:g326383; PIDN:AAA442  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 61.3%; Score 277; DB 1; Length 856;  
Best Local Similarity 94.2%; Pred. No. 4.4e-16;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLEDKWASLWMPFI 52  
1:|||||  
Db 624 NMTWMEWDREINNTSLIHSLEESQNOEKNEDELLEDKWASLWMPFI 675

RESULT 4  
VCLJLV  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)

N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03975  
R:Main-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Allzon, M.  
Cell 40, 9-17, 1985  
A:Title: Nucleotide sequence of the AIDS virus, LAV.  
A:Reference number: A90866; MUID:85099333; PMID:2981635  
A:Accession: A03975  
A:Molecule type: DNA  
A:Residues: 1-861 <MAI>  
A:Cross-references: GB:K02013; NID:g326417; PIDN:AAB59751.1; PID:g326424  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:517-861/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,146,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411  
F:616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predic

Query Match 61.3%; Score 277; DB 1; Length 861;  
Best Local Similarity 94.2%; Pred. No. 4.5e-16;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLEDKWASLWMPFI 52  
1:|||||  
Db 629 NMTWMEWDREINNTSLIHSLEESQNOEKNEDELLEDKWASLWMPFI 680

RESULT 5  
VCLJVL  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)  
N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03974  
R:Musting, M.A.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Laaky, L.A.; Capon, D.J  
Nature 313, 450-458, 1985  
A:Title: Nucleic acid structure and expression of the human AIDS lymphadenopathy retr  
A:Reference number: A93355; MUID:85111157; PMID:2982104  
A:Accession: A03974  
A:Molecule type: DNA  
A:Residues: 1-856 <MUD>  
A:Cross-references: GB:K02083; NID:g555008; PIDN:AAB59873.1; PID:g328559  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-511/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 59.3%; Score 268; DB 1; Length 856;  
Best Local Similarity 90.4%; Pred. No. 2.6e-15;  
Matches 47; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLEDKWASLWMPFI 52  
1:|||||  
Db 624 NMTWMEWDREINNTSLIHSLEESQNOEKNEDELLEDKWASLWMPFI 675

RESULT 6  
S21998  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 28  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S21998; S70425  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as dete  
A:Reference number: S21990  
A:Accession: S21998  
A:Molecule type: DNA  
A:Residues: 1-358 <STE1>  
A:Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43630.1; PID:g60183  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
Aids Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70425  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-222, 'X', 224-358 <STE2>  
A:Cross-references: EMBL:X61359; NID:g60182; PIDN:CAA43630.1; PID:g60183  
C:Superfamily: type E retrovirus env polyprotein

Query Match 57.7%; Score 261; DB 2; Length 358;  
Best Local Similarity 60.0%; Pred. No. 3.8e-15;  
Matches 51; Conservative 11; Mismatches 11; Indels 12; Gaps 1;

OY 1 NHTTWLEMDREINNTSLIHSLEESQNOEKNEDELLEDKWASLWMPFIKIQIDK 60  
1:|||||  
Db 126 NMTWMEWDREINNTSLIHSLEESQNOEKNEDELLEDKWASLWMPFI----- 175

OY 61 IEELSKTYHIEENFARIKTLGER 85  
1:|||||  
Db 176 --DITQWLMYIKIFIMYGLIGLR 198

RESULT 7  
VCLJ3W  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)



Best Local Similarity - 82.7%; Pred. NO. 1.6e-13;  
Matches 43; Conservative 7; Mismatches 2; Indels 0; Gaps 0;

**OY**    1 NHHTWLEWDREINNYTSLIHSLEEESQNOQEKEBQELLELDKASLMMNFI 52  
       | : |||::|||::| : ::|||::|||::|||::|||::|||::|||::|||::|||:  
**D6**    615 NNMTWMWEEREDINYTSEIYTLLIESQNQOEKNBQLLELDKASLMWNFDI 666

**RESULT 12**  
T12016

Query Match	54.4%	Score 246;	DB 2;	Length 852;
Best Local Similarity	57.6%	Pred. No. 2e-13;		
Matches 49;	Conservative 12;	Mismatches 12;	Indels 12;	Gaps 1;

```
QY      1 NHHTWLEMDREINNTSLIHSIIIEESQNGQEKNQEQLLELDKWSLWNFWNIKIQTIEDK 60
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      620 NNMTWMWEREIDNYSLSLIEESQNGQVKNEGELLELDKWSLSWSE----- 669
```

```
QY      61 IEELISKIYHENEIARIKKLIGER 85
      :| ::|: | : ||| |
Db      670 --DITKMLWIKIFIMVGLIGLR 692
```

RESULT 13  
S21996  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)  
C:Species: human immunodeficiency virus type 1 HIV-1  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S70422, S21996  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebrospinal fluid from patients with AIDS  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70422

Query Match	54.2%	Score	245	DB 2:	Length	357	
Best Local Similarity	56.5%	Pred. No.	8	9e-14			
Matches	48	Conservative	13	Mismatches	12	Gaps	1

```
QY      1 NHTTWLEMRREINNTSLIHSEESQNQEKEQELLEDKDWASLWNPFNRIKIQIEDK 60
          ::::: ||| :||:|||||:
Db     125 DNMTWEWEREIEINTGLIYTLIEESQNQEKEQELLEDKDWASLWNPFS----- 179
```

```
QY      61 IEELISKIYHIENEIARIKKLIGER 85
      | : ::|: | : ||| |
Db      176 --ITNMLWYIKIFIMIVGLGLGR 197
```

## RESULT 14

S22002  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 3L  
Date: 20-Feb-1995 #sequence revision 20-Feb-1995 #text change 01-Dec-2000

Query Match	54.0%;	Score 244;	DB 2;	length 358;
Best Local Similarity	80.8%;	Pred. No. 1,1e-13;		
Matches 42;	Conservative	8;	Mismatches 2;	Indels 0;
			Gaps	0;

DQ  
1 NHTTWLEWMDREINNYTSLIHSLIEESQNQQEKNCEGELLELDKWSLMMWFNI 52  
::|:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:|||||||:  
126 DNMTWMEMWDREIHNYSLTLYTLIEESQNQQEKNCEGELLELDKWSLMMWFNI 177

```

RESULT 15
S22000
envelope protein gp120/gp41 - human immunodeficiency virus type 1
C/Species: human immunodeficiency virus type 1, HIV-1
C/Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 12-Apr-1995
C/Accession: S22000
R/Steuiler, H.: Scorch-Hagenlocher, B.; Wildemann, B.; Hacke, W.
submitted to the EMBL Data Library, July 1991
A/Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as detected
A/Reference number: S21990
A/Accession: S22000
A/Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <STEX>
A/Cross-references: EMBL:X61351
A/Superfamily: type E retrovirus env polyprotein

```

Query Match Similarity	54.0%	Score 244;	DB 2;	Length 358;
Best Local Similarity	80.8%;	Pred. 1.1e-13;		
Matches 42; Conservative	8;	Mismatches 2;	Indels 0;	Gaps 0;

DQ 1 NHTTWLEWDRELNNTYSLIHLSLEESQNQQEKNDELLELDKWSLMMWNI 52  
::||:|||||:|||||:|||||:|||||:|||||:|||||:  
Dk 126 DNMTWMEDRELHNYTSLIYTLLIESQNQQEKNQDLLELDKWSLMMWFESI 177

```
Search completed: March 18, 2003, 14:37:18
Job time : 28.2167 secs
```



```

FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

```

Query Match 64.2%; Score 290; DB 1; Length 856;  
 Best Local Similarity 98.1%; Pred. No. 3,7e-18;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNOELLELDKMSLWMEFI 52
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 624 NHTTWLEMDREINNYTSLIHSLEESONQOEKNOELLELDKMSLWMEFI 675

```

```

RESULT 2
ENV_HV1H3 STANDARD; PRT; 856 AA.
AC P04624;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11707;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MDL:LINE=85228248; Pubmed=2988795;
RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
  Shaw G.M., Wong-Staal F., Reddy E.P.;
RT "HIV-III env gene products synthesized in E. coli are recognized by
  antibodies present in the sera of AIDS patients."
RL Cell 41:979-986(1985).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: M14100; AAA44679.1; -
DR HIV; M14100; ENVSHX3.
DR InterPro: IPR000328; Env-GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
  Signal.
FT SIGNAL 1 30
FT CHAIN 1 511
FT DISULFID 512 856 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT DISULFID 119 205 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 138 147 BY SIMILARITY.
FT DISULFID 218 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.
FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97188 MW; 3373C68BB84C1AFC CRC64;

```

Query Match 64.2%; Score 290; DB 1; Length 856;  
 Best Local Similarity 98.1%; Pred. No. 3,7e-18;  
 Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNOELLELDKMSLWMEFI 52
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 624 NHTTWLEMDREINNYTSLIHSLEESONQOEKNOELLELDKMSLWMEFI 675

```

```

RESULT 3
ENV_HV1HW STANDARD; PRT; 856 AA.
AC Q70626;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (LM12.3 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```



OX NCBI\_TaxID=82834;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95127297; PubMed=7826699;  
 RA Reitz G.M., Kong L., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,  
 RA Shaw G.M., Kong L., Weiss S.H., Waters D., Gallo R.C., Blattner W.;  
 RT "Viral variability and serum antibody response in a laboratory worker  
 RT infected with HIV type 1 (HIV type 1IB).";  
 RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U12055; AAA76690.1; -  
 DR GLYCSUITE: 070626; -  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00517; GP41; 1.  
 DR Pfam: PF00517; GP120; 1.  
 KW Aids; Coat protein; Polypeptide; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 511  
 FT CHAIN 512 856  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 445  
 FT DISULFID 385 418  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 289 289  
 FT CARBOHYD 295 295  
 FT CARBOHYD 301 301  
 FT CARBOHYD 332 332  
 FT CARBOHYD 339 339  
 FT CARBOHYD 356 356  
 FT CARBOHYD 386 386  
 FT CARBOHYD 392 392  
 FT CARBOHYD 397 397  
 FT CARBOHYD 406 406  
 FT CARBOHYD 448 448  
 FT CARBOHYD 463 463  
 FT CARBOHYD 611 611  
 FT CARBOHYD 616 616  
 FT CARBOHYD 624 624  
 FT CARBOHYD 637 637  
 FT CARBOHYD 674 674  
 FT CARBOHYD 750 750  
 FT CARBOHYD 816 816  
 SQ SEQUENCE 856 AA; 96938 MM; 0C2A1332CFE6687 CRC64;  
 Query Match 64.28; Score 290; DB 1; Length 856;  
 Best Local Similarity 98.1%; Pred. No. 3.7e-18;

Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 NHTTWLEMDREINNYTSLHSLEESONOEKXNDELLEDKWASLWNWNI 52  
 Db 624 NHTTWLEMDREINNYTSLHSLEESONOEKXNDELLEDKWASLWNWNI 675  
 RESULT 4  
 ENV\_HV1B8 STANDARD; PRT; 851 AA.  
 AC P04582;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 13-AUG-1987 (Rel. 05, Last sequence update)  
 DT 15-JUN-1999 (Rel. 38, Last annotation update)  
 DE Envelope glycoprotein GP160 precursor [contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (BH8 isolate) (HIV-1).  
 OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11684;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=85111123; PubMed=2578615;  
 RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,  
 RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,  
 RA Baumeister K., Ivanoff L., Petteway S.R., Pearson M.L.,  
 RA Lautenberger J.A., Papas T.S., Grayeb J., Chang N.T., Gallo R.C.,  
 RA Wong-Staal F.;  
 RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";  
 RL Nature 313:277-284(1985).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: K02011; AAA44661.1; -  
 DR HIV; K02011; ENV5B8.  
 DR GLYCSUITE: P04582; -  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW Aids; Coat protein; Polypeptide; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 506  
 FT CHAIN 507 851  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 295 295  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 440  
 FT CARBOHYD 378 440  
 FT DISULFID 385 413  
 FT CARBOHYD 88 88  
 FT CARBO

```

FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 851 AA; 96644 MM; D16A3C90857785F1 CRC64;

```

```

Query Match 61.3%; Score 277; DB 1; Length 851;
Best Local Similarity 94.2%; Pred. No. 5e-17;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 NHTTWLEMDREINNTSLIHSIESQNOEKNEDELLELDKWSLMMWNI 52
DB 619 NNTWMEWDREINNTSLIHSIESQNOEKNEDELLELDKWSLMMWNI 670

```

## RESULT 5

```

ENV_HV1B1 STANDARD; PRT; 856 AA.

```

```

AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RP MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
Josephs S.F., Doran E.R., Ratski J.A., Whitehorn E.A.,
Baumhofer K., Ivanoff L., Petteway S.R., Jr., Pearson M.L.,
Latenberger J.A., Pappas T.S., Ghayeb J., Chang N.T., Gallo R.C.,
Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III."
RL Nature 313:277-284(1985).

```

```

RN [2]
RP DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RX MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
potential glycosylation sites of the type 1 recombinant human
immunodeficiency virus envelope glycoprotein (gp120) expressed in
Chinese hamster ovary cells."
RL J. Biol. Chem. 265:10373-10382(1990).

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: M15654; AAA44205.1;
DR PIR: A03973; VCLJH3.
DR HIV: M15654; ENVSBH102.
DR InterPro: IPR000328; Env_GP41.

```

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
Signal.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT DISULFID 119 205 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .)
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .)
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .)
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .)
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .)
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .)
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .)
SO SEQUENCE 856 AA; 97224 MM; 0BFFB1A18931BB27 CRC64;

```

```

Query Match 61.3%; Score 277; DB 1; Length 856;
Best Local Similarity 94.2%; Pred. No. 5e-17;
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 NHTTWLEMDREINNTSLIHSIESQNOEKNEDELLELDKWSLMMWNI 52
DB 624 NNTWMEWDREINNTSLIHSIESQNOEKNEDELLELDKWSLMMWNI 675

```

## RESULT 6

```

ENV_HV1B1 STANDARD; PRT; 861 AA.

```

```

AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]

```



FT CARBOHD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 853 AA; 96912 MW; 3377893B6F22A8A CRC64;  
Query Match 60.6%; Score 274; DB 1; Length 853;  
Best Local Similarity 92.3%; Pred. No. 9, 2e-17;  
Matches 48; Conservative 3; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NHTTWLEMDREINNTSLHSLSIESONQOEKNEOELLELDKMSLWMMFNI 52  
1:|||||  
DB 622 NMTWMEMDREINNTSLHSLSIESONQOEKNEOELLELDKMSLWMMFNI 673  
1:|||||  
RESULT 8  
ENV\_HV1PV STANDARD; PRT; 856 AA.  
AC P03376;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=11700;  
RN [1]  
RA MEDLINE=8511157; PubMed=2982104;  
RA Mesling M.A., Smuth D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,  
RA Capon D.J.;  
RT "Nucleic acid structure and expression of the human  
RT AIDS/Lymphadenopathy retrovirus.";  
RL Nature 313:450-458(1985).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; K02083; AB559873.1; -  
DR EMBL; X01762; CAA25903.1; ALT\_SEQ.  
DR PIR; A03974; VCLJVL.  
DR HIV; K02083; ENSPV22.  
DR InterPro: IPR000328; Env GP41.  
DR InterPro: IPR000777; GP120.  
DR Pfam; PF00516; GP120; 1.  
DR Pfam; PF00517; GP41; 1.  
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;  
KW Signal.  
FT SIGNAL 1 30  
FT CHAIN 31 511 EXTERIOR MEMBRANE GLYCOPROTEIN.  
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.  
FT DISULFID 54 74 BY SIMILARITY.  
FT DISULFID 119 205 BY SIMILARITY.  
FT DISULFID 126 196 BY SIMILARITY.  
FT DISULFID 131 157 BY SIMILARITY.  
FT DISULFID 218 247 BY SIMILARITY.  
FT DISULFID 228 239 BY SIMILARITY.

FT DISULFID 296 331 BY SIMILARITY.  
FT DISULFID 378 445 BY SIMILARITY.  
FT DISULFID 385 418 BY SIMILARITY.  
FT CARBOHD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 856 AA; 97339 MW; 5FCDB1DC3C120993 CRC64;  
Query Match 59.3%; Score 268; DB 1; Length 856;  
Best Local Similarity 90.4%; Pred. No. 3, 1e-16;  
Matches 47; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
QY 1 NHTTWLEMDREINNTSLHSLSIESONQOEKNEOELLELDKMSLWMMFNI 52  
1:|||||  
DB 624 NMTWMEMDREINNTSLHSLSIESONQOEKNEOELLELDKMSLWMMFNI 675  
1:|||||  
RESULT 9  
ENV\_HV1W1 STANDARD; PRT; 856 AA.  
AC P31872;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane  
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
GN ENV.  
OS Human immunodeficiency virus type 1 (WMJ1 isolate) (HIV-1).  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_Taxid=31678;  
RN [1]  
RA MEDLINE=86218077; PubMed=2423250;  
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,  
RA Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;  
RT "Identification and characterization of conserved and variable  
RT regions in the envelope gene of HTLV-III/LAV, the retrovirus of  
RT AIDS.";  
RL Cell 45:637-648(1986).  
CC -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM  
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO  
CC WAS PERINATALLY INFECTED BY HER MOTHER.  
DR PIR: A24774; VCLJW.  
DR InterPro: IPR000328; Env GP41.  
DR InterPro: IPR000777; GP120.



[illegible]

FT	CARBOHYD	602	602	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD <td>607<td>607<td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td></td></td>	607 <td>607<td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td></td>	607 <td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td>	N-LINKED (GLCNAC. . .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD <td>616<td>616<td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td></td></td>	616 <td>616<td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td></td>	616 <td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td>	N-LINKED (GLCNAC. . .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD <td>628<td>628<td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td></td></td>	628 <td>628<td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td></td>	628 <td>N-LINKED (GLCNAC. . .)<th>(POTENTIAL).</th> </td>	N-LINKED (GLCNAC. . .) <th>(POTENTIAL).</th>	(POTENTIAL).
FT	CARBOHYD <td>647 AA:<td>96135 MW;<td>0A901317FD7FF2AB<td>CRC64;</td> </td></td></td>	647 AA: <td>96135 MW;<td>0A901317FD7FF2AB<td>CRC64;</td> </td></td>	96135 MW; <td>0A901317FD7FF2AB<td>CRC64;</td> </td>	0A901317FD7FF2AB <td>CRC64;</td>	CRC64;
SO	SEQUENCE	847 AA: <td>96135 MW;<td>0A901317FD7FF2AB<td>CRC64;</td> </td></td>	96135 MW; <td>0A901317FD7FF2AB<td>CRC64;</td> </td>	0A901317FD7FF2AB <td>CRC64;</td>	CRC64;
Query Match			55.8%;	Score 252;	DB 1: Length 847;
Best Local Similarity			58.1%;	Pred. No. 7.7e-15;	
Matches	50;	Conservative	14;	Mismatches	8; Indels 14; Gaps 2
Qy	1	NHTTLEMDREINNTYSLHSILIESQNOEKNQDELLDLKDWASLMMFNIRKIQIEDK	60		
DB	615	NNMTWMEIREIDNTYNTLYLTILESQNOEKNQDELLDLKDWASLMMFNFI	666		
Qy	61	IEELISK-IYHENEIARKIKLIGER	85		
DB	667	-----SKWLWIKRIFIMIVGLVGR	687		
RESULT 12					
ENV_HV153					
ID	ENV_HV153	STANDARD:	PRT:	852 AA.	
AC	P19549:				
DI	01-FEB-1991 (Rel. 17, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Envelope glycoprotein gp160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].				
DE	GN				
OS	Human immunodeficiency virus type 1 (SP3 isolate) (HIV-1).				
OC	Viruses; Retroid viruses; Retroviridae; Lentiviruts.				
OX	NCBI_TaxID=11690;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90317906; PubMed=2370688;				
RA	York-Higgins D., Cheng-Mayer C., Bauer D., Levy J.A., Dina D.;				
RT	"Human immunodeficiency virus type 1 cellular host range, replication, and cytopathicity are linked to the envelope region of the viral genome.";				
RL	J. Virol. 64:4016-4020(1990).				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use. By non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement/ or send an email to license@isb-sib.ch">http://www.isb-sib.ch/announcement/</a> or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL: M38427; AAA5067.1; -				
DR	HIV; M38427; ENVSP33.				
DR	InterPro; IPR003128; ENV_GP41.				
DR	InterPro; IPR000777; GP120.				
DR	Pfam; PF00516; GP120; 1.				
DR	Pfam; PF00517; GP41; 1.				
KW	AIDS; Coat protein; Glycoprotein; Transmembrane; Signal.				
KW	SIGNAL.				
FT	SIGNAL	1	31	BY SIMILARITY.	
FT	CHAIN	32	506	EXTERIOR MEMBRANE GLYCOPROTEIN.	
FT	CHAIN	507	852	TRANSMEMBRANE GLYCOPROTEIN.	
FT	DISULFID	53	73	BY SIMILARITY.	
FT	DISULFID	118	206	BY SIMILARITY.	
FT	DISULFID	125	197	BY SIMILARITY.	
FT	DISULFID	130	156	BY SIMILARITY.	
FT	DISULFID	219	248	BY SIMILARITY.	
FT	DISULFID	229	240	BY SIMILARITY.	
FT	DISULFID	297	331	BY SIMILARITY.	
FT	DISULFID	377	439	BY SIMILARITY.	
FT	DISULFID	397	412	BY SIMILARITY.	
FT	DISULFID	384	412	BY SIMILARITY.	
FT	CARBOHYD	87	87	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	129	129	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	141	141	N-LINKED (GLCNAC. . .)	(POTENTIAL).

```

FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 155 155 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 159 159 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 242 242 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 263 263 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 385 385 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 391 391 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 405 405 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 96663 MW; EETBBFBD23C9910D CRC64;

```

```

Query Match 55.8%; Score 252; DB 1; Length 852;
Best Local Similarity 82.7%; Pred. No. 7, 7e-15;
Matches 43; Conservative 8; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 NHTTWLEMPREINNTYSLIHSILEESONOQKNEQLELDKWSLMMNFI 52
DB 620 NMTWMEWEREDNTYSLIYTLLEESONOQKNEQLELDKWSLMMNFI 671

```

```

RESULT 13
ENV_HV1M2 STANDARD; PRT; 847 AA.
AC P05880;
DT 01-NOV-1988 (rel. 09, Created)
DT 01-NOV-1988 (rel. 09, Last sequence update)
DT 15-JUL-1999 (rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV
OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11705;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86235450; PubMed=3012778;
RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,
Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;
RT "Genetic variation in HIV-III/LAV over time in patients with AIDS or
at risk for AIDS."
RT Science 233:1548-1553(1986).
RL -1- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM
BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
WAS PERINATALLY INFECTED BY HER MOTHER.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

DR EMBL: M12507; AAB12990.1; -
DR HIV: M12307; ENV$WMJ2.
DR InterPro: IPR000328; ENV-GP41.
DR InterPro: IPR000777; GP120.

```

```

DR Pfam: PF00516; GP120; 1.
KW Pfam: PF00517; GP41; 1.
DR AIDS: Coat protein; Glycoprotein; Transmembrane;
KW Signal.
FT CHAIN 1 29
FT CHAIN 501 501
FT CHAIN 502 847
FT DISULFID 53 73
FT DISULFID 118 202
FT DISULFID 125 193
FT DISULFID 130 152
FT DISULFID 215 244
FT DISULFID 225 236
FT DISULFID 293 326
FT DISULFID 372 435
FT DISULFID 379 408
FT CARBOHYD 87 87
FT CARBOHYD 134 134
FT CARBOHYD 140 140
FT CARBOHYD 151 151
FT CARBOHYD 155 155
FT CARBOHYD 183 183
FT CARBOHYD 184 184
FT CARBOHYD 194 194
FT CARBOHYD 231 231
FT CARBOHYD 238 238
FT CARBOHYD 259 259
FT CARBOHYD 273 273
FT CARBOHYD 286 286
FT CARBOHYD 292 292
FT CARBOHYD 327 327
FT CARBOHYD 334 334
FT CARBOHYD 350 350
FT CARBOHYD 356 356
FT CARBOHYD 380 380
FT CARBOHYD 386 386
FT CARBOHYD 390 390
FT CARBOHYD 400 400
FT CARBOHYD 438 438
FT CARBOHYD 450 450
FT CARBOHYD 602 602
FT CARBOHYD 607 607
FT CARBOHYD 616 616
FT CARBOHYD 628 628
SQ SEQUENCE 847 AA; 96466 MW; CDLE33073AA5B0CAE CRC64;

```

```

Query Match 54.2%; Score 245; DB 1; Length 847;
Best Local Similarity 56.5%; Pred. No. 3, 1e-14;
Matches 48; Conservative 14; Mismatches 11; Indels 12; Gaps 1;

```

```

OY 1 NHTTWLEMPREINNTYSLIHSILEESONOQKNEQLELDKWSLMMNFIKIKQIEDK 60
DB 615 DNTWMEWEREDNTYSLIYTLLEESONOQKNEQLELDKWSLMMNFI----- 664
OY 61 IEELISKIYHIENEIARIKKLIGER 85
DB 665 --DITWMLWYKIFIMIVGSLIGLR 687

```

```

RESULT 14
ENV_HV1M2 STANDARD; PRT; 856 AA.
AC P05877;
DT 01-NOV-1988 (rel. 09, Created)
DT 01-NOV-1988 (rel. 09, Last sequence update)
DT 15-JUL-1999 (rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV
OS Human immunodeficiency virus type 1 (MN isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxId=11696;
RN [1]

```

[illegible][illegible]



```
FT CARBOHYD 353 353 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 97043 MW; 849B08BCBAFF7008 CRC64;
```

```
Query Match 53.3%; Score 241; DB 1; Length 853;
Best Local Similarity 85.7%; Pred. No. 7.1e-14;
Matches 42; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
```

```
OY 4 TWLEMDREINNYTSLHSLIEESONOEKNEOELLELDKXWASLWNNFNI 52
DB 624 TWMEEREINDNYTGILIRLIEESOTOOEKNEOELLELDKXWASLWNNFNI 672
```

Search completed: March 18, 2003, 14:34:02  
Job time : 13.1429 secs

**THIS PAGE BLANK (OPTO)**

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:01 ; Search time 47.734 seconds  
(without alignments)  
366.908 Million cell updates/sec

Title: US-09-877-606-4  
Perfect score: 452  
Sequence: 1 NHTWLEMDREINNYSLIH.....SKYHIENELARIKLIGER 85

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP-archaea:\*  
2: SP-bacteria:\*  
3: SP-fungi:\*  
4: SP-human:\*  
5: SP-invertebrate:\*  
6: SP-mammal:\*  
7: SP-mhc:\*  
8: SP-organelle:\*  
9: SP-phage:\*  
10: SP-plant:\*  
11: SP-rodent:\*  
12: SP-virus:\*  
13: SP-vertebrate:\*  
14: SP-unclassified:\*  
15: SP-virus:\*  
16: SP-bacteriap:\*  
17: SP-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	290	64.2	748	15	Q70606 human immun
2	290	64.2	752	15	Q70604 human immun
3	290	64.2	752	15	Q70605 human immun
4	286	63.3	747	15	Q70607 human immun
5	286	63.3	752	15	Q70608 human immun
6	285	63.1	856	15	Q90SM7 human immun
7	282	62.4	852	15	Q89797 human immun
8	280	61.9	856	15	Q92877 simian-huma
9	277	61.3	645	15	Q993A6 human immun
10	277	61.3	851	15	Q782A3 human immun
11	277	61.3	854	15	Q85582 human immun
12	277	61.3	854	15	Q72502 human immun
13	277	61.3	856	15	Q74599 human immun
14	277	61.3	856	15	Q74090 human immun
15	273	60.4	854	15	Q90178 human immun
16	273	60.4	854	15	Q78705 human immun

17	270	59.7	757	15	Q9q722 human immun
18	265	58.6	851	15	Q56110 human immun
19	265	58.6	851	15	Q9e1r7 human immun
20	264	58.4	848	15	Q69990 human immun
21	262	58.0	727	15	Q9q723 human immun
22	261	57.7	358	15	Q78120 human immun
23	261	57.7	616	15	Q993B0 human immun
24	261	57.7	618	15	Q993B2 human immun
25	261	57.7	854	15	Q78225 human immun
26	260	57.5	838	15	Q03806 human immun
27	260	57.5	848	15	Q69988 human immun
28	260	57.5	855	15	Q03805 human immun
29	260	57.5	862	15	Q9E1S2 human immun
30	259	57.3	635	15	Q90U82 human immun
31	259	57.3	852	15	Q69992 human immun
32	257	56.9	856	15	Q72993 human immun
33	257	56.9	859	15	Q6Q863 human immun
34	257	56.9	863	15	Q9wJu4 human immun
35	257	56.9	864	15	Q9yP48 human immun
36	256	56.6	851	15	Q8Q852 human immun
37	256	56.6	854	15	Q56112 human immun
38	256	56.6	863	15	Q9wJu6 human immun
39	256	56.6	863	15	Q42031 human immun
40	256	56.6	864	15	Q9wJu1 human immun
41	255	56.4	847	15	Q69996 human immun
42	255	56.4	858	15	Q71974 human immun
43	254	56.2	849	15	Q73368 human immun
44	254	56.2	858	15	Q8Q867 human immun
45	254	56.2	858	15	Q8Q865 human immun

## ALIGNMENTS

RESULT 1	Q70606	PRELIMINARY;	PRT;	748	AA.
ID	Q70606				
AC	Q70606				
DT	01-NOV-1996 (TREMBLrel. 01, Created)				
DT	01-NOV-1996 (TREMBLrel. 01, Last sequence update)				
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)				
DE	Envelope glycoprotein (Fragment).				
ENV.					
OS	Human immunodeficiency virus type 1.				
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.				
OX	NCBI_TaxID-11676;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-LW881;				
RX	MEDLINE-95127297; PubMed-7826699;				
RA	Shaw G., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,				
RA	Reitz M., Weiss S., Waters D., Gallo R., Blattner W.;				
RT	"Viral variability and serum antibody response in a laboratory worker				
RT	infected with HIV type 1 (HIV type IIIB).";				
RL	AIDS Res. Hum. Retroviruses 10:1143-1155(1994).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-LW881;				
RA	Mulder K.E.;				
RL	Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.				
DR	EMBL; U12032; AAA76668.1;				
DR	InterPro: IPR000328; Env.GP41.				
DR	InterPro: IPR000777; GP120.				
DR	Pfam: PF00516; GP120; 1.				
KW	Pfam: PF00517; GP41; 1.				
FT	AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.				
FT	NON_TER				
SO	SEQUENCE	748	748		
		748	748		
		84224	MM;	56BEDF186C67694B	CRC64;
Query Match		64.2%	Score 290;	DB 15;	Length 748;
Best Local Similarity		98.1%;	Pred. No. 4.5e-16;		
Matches	51;	Conservative	1;	Mismatches	0;
				Indels	0;
				Gaps	0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKASLWVFN 52  
DB 620 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKASLWVFN 671

## RESULT 2

ID Q70604 PRELIMINARY; PRT; 752 AA.  
AC Q70604;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LM851;  
RX MEDLINE=95127297; Pubmed=7826699;  
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,  
Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;  
RT "Viral variability and serum antibody response in a laboratory worker  
infected with HIV type 1 (HIV type IIIB).";  
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LM851;  
RA Mulder K.E.;  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U12030; AAA7666.1; -;  
DR InterPro: IPR000328; Env\_Gp41.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KM AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.  
FT NON\_TER 752  
SQ SEQUENCE 752 AA; 84894 MW; 8B30AE894013B45A CRC64;

## Query Match

Best Local Similarity 64.2%; Score 290; DB 15; Length 752;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKASLWVFN 52  
DB 624 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKASLWVFN 675

## RESULT 3

ID Q70605 PRELIMINARY; PRT; 752 AA.  
AC Q70605;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LM852;  
RX MEDLINE=95127297; Pubmed=7826699;  
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,  
Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;  
RT "Viral variability and serum antibody response in a laboratory worker  
infected with HIV type 1 (HIV type IIIB).";  
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LM852;

RA Mulder K.E.;  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U12031; AAA7667.1; -;  
DR InterPro: IPR000328; Env\_Gp41.  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KM AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.  
FT NON\_TER 752  
SQ SEQUENCE 752 AA; 84894 MW; 8B30AE894013B45A CRC64;

## Query Match

Best Local Similarity 64.2%; Score 290; DB 15; Length 752;  
Matches 51; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKASLWVFN 52  
DB 624 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKASLWVFN 675

## RESULT 4

ID Q70607 PRELIMINARY; PRT; 747 AA.  
AC Q70607;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LM87-1;  
RX MEDLINE=95127297; Pubmed=7826699;  
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,  
Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;  
RT "Viral variability and serum antibody response in a laboratory worker  
infected with HIV type 1 (HIV type IIIB).";  
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LM87-1;  
RA Mulder K.E.;  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U12034; AAA7669.1; -;  
DR InterPro: IPR000328; Env\_Gp41.  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KM AIDS; Coat protein; Glycoprotein; Polypeptide; Transmembrane.  
FT NON\_TER 747  
SQ SEQUENCE 747 AA; 84250 MW; 732E836A52245F14 CRC64;

## Query Match

Best Local Similarity 63.3%; Score 286; DB 15; Length 747;  
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKASLWVFN 52  
DB 619 NHTTWLEMDREINNYTSLIHSLEESONQOEKNEQELLELDKASLWVFN 670

## RESULT 5

ID Q70608 PRELIMINARY; PRT; 752 AA.  
AC Q70608;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.

```

OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LW87-2;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
  Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
  "Viral variability and serum antibody response in a laboratory worker
  infected with HIV type 1 (HIV type IIIB).";
RT AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LW87-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U12053; AAA76670.1; -.
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KM AIDS: Coat protein; Glycoprotein; Polyprotein; Transmembrane.
FT NON_TER
SQ SEQUENCE 752 AA; 84780 MW; 708672A2D0C0E8F8 CRC64;

Query Match          63.3%; Score 286; DB 15; Length 752;
Best Local Similarity 98.1%; Pred. No. 9.6e-16;
Matches 51; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWEMDEINNYTSLIHSLEESONQOEKNEQELLELDKASLWNWNI 52
Db 624 NHTTWEMDEINNYTSLIHSLEESONQOEKNEQELLELDKASLWNWNI 675

RESULT 6
Q90SM7 PRELIMINARY; PRT; 856 AA.
AC Q90SM7;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-HXB2;
RA Altman-Onal V., Cheynet V., Verrier B.;
  "Mutations and transcriptional alterations associated with the
  RT downregulation of HIV-1 envelope glycoprotein expression following
  acute cytopathic effects.";
RT acute cytopathic effects.";
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF358141; AAK49977.1; -.
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KM AIDS: Coat protein; Glycoprotein; Polyprotein; Transmembrane.
SQ SEQUENCE 856 AA; 97126 MW; 9458D02B2FD734B3 CRC64;

Query Match          63.1%; Score 285; DB 15; Length 856;
Best Local Similarity 96.2%; Pred. No. 1.3e-15;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWEMDEINNYTSLIHSLEESONQOEKNEQELLELDKASLWNWNI 52
Db 624 NHTTWEMDEINNYTSLIHSLEESONQOEKNEQELLELDKASLWNWNI 675

RESULT 7

```

```

Q89797
ID Q89797 PRELIMINARY; PRT; 852 AA.
AC Q89797;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LW90-2;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
  Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
  "Viral variability and serum antibody response in a laboratory worker
  infected with HIV type 1 (HIV type IIIB).";
RT AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-LW90-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U12053; AAA76685.1; -.
DR EMBL: U12036; AAA76671.1; -.
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KM AIDS: Coat protein; Envelope protein; Glycoprotein; Polyprotein;
  Transmembrane.
SQ SEQUENCE 852 AA; 96347 MW; 4E86522D4EB330F CRC64;

Query Match          62.4%; Score 282; DB 15; Length 852;
Best Local Similarity 96.2%; Pred. No. 2.3e-15;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWEMDEINNYTSLIHSLEESONQOEKNEQELLELDKASLWNWNI 52
Db 620 NHTTWEMDEINNYTSLIHSLEESONQOEKNEQELLELDKASLWNWNI 671

RESULT 8
Q92877 PRELIMINARY; PRT; 856 AA.
AC Q92877;
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Simian-Human immunodeficiency virus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_Taxid=57667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9908984; PubMed=9882298;
RA Cayabyab M., Karlsson G.B., Etemad-Moghadam B.A., Hofmann W.,
  Steenbeke T., Halloran M., Fanton J.W., Athelms M.K., Letvin N.L.,
  Sodroski J.G.;
  "Changes in human immunodeficiency virus type 1 envelope glycoproteins
  responsible for the pathogenicity of a multiply passaged simian-human
  immunodeficiency virus (SHIV-HXB2).";
RT responsible for the pathogenicity of a multiply passaged simian-human
  immunodeficiency virus (SHIV-HXB2).";
RL J. Virol. 73:976-984(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Cayabyab M., Karlsson G.B., Etemad-Moghadam B., Hofmann W.,
  Halloran M., Athelms M.W., Letvin N.L., Sodroski J.G.;
  Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF041850; AAD12142.1; -.
DR InterPro: IPR000328; Env_GP41.

```

DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SEQUENCE 856 AA; 97151 MW; C50BE0388FB73659 CRC64;  
Query Match 61.9%; Score 280; DB 15; Length 856;  
Best Local Similarity 63.5%; Pred. No. 3.4e-15;  
Matches 54; Conservative 9; Mismatches 10; Indels 12; Gaps 1;  
QY 1 NHTTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFNKIQIEDK 60  
DB 624 NHTTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFN----- 673  
QY 61 IEETISKYHIEHETARIKIKLIGER 85  
DB 674 --DITNMLMYIKLFIMIVGLVGLR 696  
RESULT 9  
ID 0993A6 PRELIMINARY; PRT; 645 AA.  
AC 0993A6;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE Truncated envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=1007;  
RC MEDLINE=21192672; PubMed=11287644;  
RA Suman S., Lockey T.D., Slobod K.S., Jones B., Riberd J.M.,  
RA White S.W., Doherty P.C., Hurtwitz J.L.;  
RT "Localization of CD4+ T cell epitope hotspots to exposed strands of  
RT HIV envelope glycoprotein suggests structural influences on antigen  
RT processing.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4587-4592(2001).  
DR EMBL: AF321563; AAK18810.1; -;  
DR InterPro: IPR000328; Env-GP41.  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KM AIDS; Coat protein; Envelope protein; Glycoprotein; Polyprotein;  
KW Transmembrane.  
FT NON TER 1  
SQ SEQUENCE 645 AA; 72485 MW; B076514BE93362EC CRC64;  
Query Match 61.3%; Score 277; DB 15; Length 645;  
Best Local Similarity 94.2%; Pred. No. 4.5e-15;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NHTTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFN 52  
DB 593 NHTTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFN 644  
RESULT 10  
ID 078243 PRELIMINARY; PRT; 851 AA.  
AC 078243;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE Env Polyprotein.  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]

RP SEQUENCE FROM N.A.  
RA Carlini F., Federic M., Equestre M., Ricci S., Ratti G., Zibai O.,  
RA Verani P., Rossi G.B.;  
RT "Sequence analysis of HIV-1 proviral DNA from a non producer  
RT chronically infected HUT-78 cellular clone.";  
RL J. Viral Diseases 1:40-55(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89352106; PubMed=2765297;  
RA Federico M., Tifti F., Butto S., Orecchia A., Carlini F., Taddeo B.,  
RA Macchi B., Mangiano N., Verani P., Rossi G.;  
RT "Biological and molecular characterization of producer and non  
RT producer clones from HUT-78 infected with a patient HIV isolate.";  
RL AIDS Res. Hum. Retroviruses 5:385-396(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Tifti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,  
RA Borsetti A., Saggio I., Verani P., Rossi G.;  
RT "Variability of HIV-1 virus: characteristics of an infected but not  
RT productive clone.";  
RL Int. J. Immunopharmacol. 3:17-23(1990).  
DR EMBL: Z11530; CAA77628.1; -;  
DR InterPro: IPR000328; Env-GP41.  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KM AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
SQ SEQUENCE 851 AA; 96630 MW; 1A37678B97E98027 CRC64;  
Query Match 61.3%; Score 277; DB 15; Length 851;  
Best Local Similarity 94.2%; Pred. No. 5.9e-15;  
Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 1 NHTTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFN 52  
DB 619 NHTTWLEMDREINNTSLIHSIESQNOQEKNEQELLELDKMSLWMMFN 670  
RESULT 11  
ID 085582 PRELIMINARY; PRT; 854 AA.  
AC 085582;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE Envelope polyprotein.  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=86281827; PubMed=3016298;  
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,  
RA Martin M.A.;  
RT "Production of acquired immunodeficiency syndrome-associated  
RT retrovirus in human and nonhuman cells transfected with an infectious  
RT molecular clone.";  
RL J. Virol. 59:284-291(1986).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;  
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Buckler C.E.;  
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92219406; PubMed=1373204;  
RA Dai L.C., Litnau R., Takahashi K., Ennis F.A.;  
RT "Mutation of human immunodeficiency virus type 1 at amino acid 585 on  
RT gp41 results in loss of killing by CD8+ A24-restricted cytotoxic T

RT Lymphocytes.";  
 RL J. Virol. 66:3151-3154(1992).  
 DR EMBL: M19921; AAA44992.1; -.  
 DR InterPro: IPR000328; Env\_GpA1.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120. 1.  
 DR Pfam: PF00517; GP41. 1.  
 KW AIDS; Coat protein; glycoprotein; polyprotein; Transmembrane.  
 SQ SEQUENCE 854 AA; 97124 MW; ABA0A1BF36D5555A CRC64;

Query Match 61.3%; Score 277; DB 15; Length 854;  
 Best Local Similarity 94.2%; Pred. No. 6e-15;  
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLHSLIESONOQEKNEQELLELDKWSLMMNFI 52  
 Db 622 NNTWMEDEINNTSLHSLIESONOQEKNEQELLELDKWSLMMNFI 673

## RESULT 12

ID 072502 PRELIMINARY; PRT; 854 AA.  
 AC 072502;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ENV polyprotein.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NL4-3;  
 RX MEDLINE=96036482; PubMed=7483282;  
 RA Salzman M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,  
 RA Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;  
 RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes  
 from primary virus cultures using the polymerase chain reaction.";  
 RL Virology 213:80-86(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.

RC STRAIN=NL4-3;  
 RX MEDLINE=86281827; PubMed=3016298;  
 RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,  
 RA Martin M.A.;  
 RT "Production of acquired immunodeficiency syndrome-associated  
 retrovirus in human and nonhuman cells transfected with an infectious  
 RT molecular clone.";  
 RL J. Virol. 59:284-291(1986).  
 DR EMBL: U26942; AAB60578.1; -.  
 DR InterPro: IPR000328; Env\_GpA1.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120. 1.  
 DR Pfam: PF00517; GP41. 1.  
 KW AIDS; Coat protein; glycoprotein; polyprotein; Transmembrane.  
 FT CONFLICT 214 214 H -> L (IN REF. 2).  
 FT CONFLICT 530 530 A -> S (IN REF. 2).  
 FT CONFLICT 739 739 G -> D (IN REF. 2).  
 SQ SEQUENCE 854 AA; 97005 MW; FE2264B3841D1220 CRC64;

Query Match 61.3%; Score 277; DB 15; Length 854;  
 Best Local Similarity 94.2%; Pred. No. 6e-15;  
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLHSLIESONOQEKNEQELLELDKWSLMMNFI 52  
 Db 622 NNTWMEDEINNTSLHSLIESONOQEKNEQELLELDKWSLMMNFI 673

RESULT 13  
 ID 074599 PRELIMINARY; PRT; 856 AA.

AC 074599;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ENV.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MCX1;  
 RX MEDLINE=90101366; PubMed=1688473;  
 RA Cloyd M.W., Moore B.E.;  
 RT "Spectrum of Biological Properties of Human Immunodeficiency Virus  
 (HIV-1) Isolates.";  
 RL Virology 174:103-116(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MCX1;  
 RA Iwataki Y.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: D86058; BAA12995.1; -.  
 DR InterPro: IPR000328; Env\_GpA1.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120. 1.  
 DR Pfam: PF00517; GP41. 1.  
 KW AIDS; Coat protein; glycoprotein; polyprotein; Transmembrane.  
 SQ SEQUENCE 856 AA; 97287 MW; 238042A234C56585 CRC64;

Query Match 61.3%; Score 277; DB 15; Length 856;  
 Best Local Similarity 94.2%; Pred. No. 6e-15;  
 Matches 49; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLHSLIESONOQEKNEQELLELDKWSLMMNFI 52  
 Db 624 NNTWMEDEINNTSLHSLIESONOQEKNEQELLELDKWSLMMNFI 675

RESULT 14  
 ID 074090 PRELIMINARY; PRT; 856 AA.  
 AC 074090;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE ENV.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PM213;  
 RX MEDLINE=90101366; PubMed=1688473;  
 RA Cloyd M.W., Moore B.E.;  
 RT "Spectrum of Biological Properties of Human Immunodeficiency Virus  
 (HIV-1) Isolates.";  
 RL Virology 174:103-116(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PM213;  
 RA Iwataki Y.

RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: D86059; BAA13003.1; -.  
 DR InterPro: IPR000328; Env\_GpA1.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120. 1.  
 DR Pfam: PF00517; GP41. 1.  
 KW AIDS; Coat protein; glycoprotein; polyprotein; Transmembrane.  
 SQ SEQUENCE 856 AA; 97396 MW; FE3E784C423C108C CRC64;

Query Match 61.3%; Score 277; DB 15; Length 856;





GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: March 25, 2003, 12:52:44 ; Search time 2395.91 Seconds  
(without alignments)  
574.569 Million cell updates/sec

Title: US-09-877-606-4  
Perfect score: 452  
Sequence: 1 NHPTWLEMDREINNYTSLIH.....SKYHENEIARIKILIGER 85

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Command line parameters:  
-MODE=frame\_plus.p2n.model -DEV=ylh  
-Q/cgn2.1/USPTO-spool/US09877606/runat.18032003\_134800\_17705/app.query.fasta.1.1059  
-DB=EST -Qfmt=fastap -SUFFIX=est -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -STRAT=1 -BND=1 -MATRIX=biosum62 -TRANS=human4.0.cdi -LIST=45  
-DOCALIGN=200 -THR\_SCORE=DCT -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=2000000000  
-USER=US09877606\_gcgn.1.1.4134\_g-runat.18032003\_134800\_17705 -NCPu=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEORDER -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_etc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pin:\*  
21: em\_gss\_vrl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_tod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	136	30.1	320	14	T38676	T38676 EST104204 S
2	124	27.4	370	14	T36326	T36326 EST101246 S
3	122	27.0	323	14	T36997	T36997 EST102040 S
4	115	25.4	363	14	T36351	T36351 EST101276 S
5	114	25.2	210	14	T38672	T38672 EST104198 S
6	107	23.7	263	14	T37110	T37110 EST102177 S
7	86	19.0	747	13	BM166231	BM166231 EST568754
8	83.5	18.5	284	14	BP027004	BP027004 BP027004
9	83.5	18.5	601	17	BH280804	BH280804 CH230-71L
10	83	18.4	541	17	AQ376798	AQ376798 RPK11-15
11	83	18.4	543	17	AQ793621	AQ793621 HS_5421.B
12	82.5	18.3	406	9	AT676261	AT676261 eTerm000
13	82	18.1	344	13	BM170439	BM170439 EST572962
14	82	18.1	364	13	BM170211	BM170211 EST572734
15	82	18.1	696	13	BI151755	BI151755 602915918
16	82	18.1	915	13	BI654290	BI654290 603280225
17	82	18.1	978	17	AZ693053	AZ693053 ENTJX74TR
18	81	17.9	955	14	BQ216438	BQ216438 AGENCOURT
19	80.5	17.8	695	9	A1362015	A1362015 pat.PX007
20	80.5	17.8	866	11	AK005825	AK005825 Mus muscu
21	80	17.7	848	17	BH151419	BH151419 ENTQ152TF
22	80	17.7	975	17	CNS0620E	AL422652 T3 end of
23	79.5	17.6	473	10	AW375844	AW375844 RC3-CT019
24	79.5	17.6	515	14	BO090125	BO090125 rc61h05.Y
25	79.5	17.6	556	10	BE070210	BE070210 OVA-BP040
26	79.5	17.6	639	9	AL639656	AL639656 AL639656
27	79.5	17.6	986	12	BG748525	BG748525 602706214
28	79	17.5	340	12	BF016856	BF016856 ux17e10.Y
29	79	17.5	529	13	BM276448	BM276448 PfESToa8
30	79	17.5	874	17	BH131484	BH131484 ENTOM27TF
31	79	17.5	889	17	AZ532507	AZ532507 ENTBS54TR
32	79	17.5	904	17	BH131547	BH131547 ENTOS30TR
33	79	17.5	909	17	AZ543458	AZ543458 ENTBU13TF
34	79	17.5	934	17	BH137560	BH137560 ENTQ125TF
35	78.5	17.4	304	17	AQ137562	AQ137562 HS_2213.A
36	78.5	17.4	889	17	CNS045W5	AL275918 Tetradon
37	78	17.3	366	12	BG100679	BG100679 uy12d02.Y
38	78	17.3	461	10	AW783430	AW783430 Ta26107.Y
39	78	17.3	461	10	AW827926	AW827926 ra58h05.Y
40	78	17.3	483	17	B56487	B56487 CIT-HSP-200
41	78	17.3	485	10	AW870869	AW870869 ra51g05.Y
42	78	17.3	574	13	BG917674	BG917674 602820703
43	78	17.3	618	13	BM281579	BM281579 K117C12.Y
44	78	17.3	707	17	AQ419358	AQ419358 RPK1-11-2
45	78	17.3	841	17	AZ528712	AZ528712 ENTBG49TF

ALIGNMENTS

RESULT 1	T38676/c	320 bp	mRNA	linear	EST 11-JAN-1995
LOCUS	T38676				
DEFINITION	EST104204 S. cerevisiae strain X2180-1A Saccharomyces cerevisiae cDNA 3' end, mRNA sequence.				
ACCESSION	T38676				
VERSION	T38676.1	GI:622493			
KEYWORDS	EST.				
SOURCE	baker's yeast.				
ORGANISM	Saccharomyces cerevisiae				
	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
REFERENCE	1 (bases 1 to 320)				
AUTHORS	Weinstock, K.				
TITLE	Saccharomyces cerevisiae cDNAs				

```

JOURNAL
COMMENT      Unpublished (1995)
              Other ESTs: TC473
              Contact: Weinstock, K. and Venter, J.C.
              The Institute for Genomic Research
              932 Clopper Rd, Galthersburg, MD 20878
              Tel: 3018699056
              Fax: 3018699423
              Email: tdbinfo@tdb.tigr.org
              For clone availability please contact the TIGR Database
              (tdbinfo@tdb.tigr.org)
              Seq primer: M13-21.

FEATURES
  source
    1..320
      /organism="Saccharomyces cerevisiae"
      /strain="X2180-1A"
      /db_xref="taxon:4932"
      /clone_1lb="S. cerevisiae strain X2180-1A"
      /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI"

BASE COUNT   97 a   54 c   58 g   106 t   5 others

ORIGIN
Alignment Scores:
Pred. No.:      6.49e-07      Length:      320
Score:          136.00      Matches:      23
Percent Similarity: 96.97%      Conservative: 9
Best Local Similarity: 69.70%      Mismatches:  1
Query Match:     30.09%      Indels:      0
DB:              14      Gaps:          0

US-09-877-606-4 (1-85) x T38676 (1-320)

QY   53  LysllElySGhllEgluApSlyllEgluGlullELeSerLysllEtyrHisllEglu  72
      ::::::::::::::::::::::::::::::::::::::::::::::::::::
Db   205 ACAAAGAAACAACCTTGAAGACACAAAGGTTGACGAATTCCTTCACAAAATTATCACTTGAA  146
      |||||||
QY   73  AsngluIleAlaArgllElySlySleuIlEglYgluArg  85
      |||||||
Db    145 AATGAGGTTGCCAGATTAAAGAAATTAGTGGCGAAGCG  107
      |||||||

RESULT 2
T36326      370 bp      mRNA      linear      EST 11-JAN-1995
LOCUS       T36326/c
DEFINITION cDNA 3' end, mRNA sequence.
ACCESSION   T36326
VERSION     T36326.1
KEYWORDS    GI:620143
SOURCE      EST.
  ORGANISM  Baker's yeast.
            Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
            1 (bases 1 to 370)
            Weinstock, K.
            Saccharomyces cerevisiae cDNAs
            Unpublished (1995)
            Other ESTs: TC473
            Contact: Weinstock, K. and Venter, J.C.
            The Institute for Genomic Research
            932 Clopper Rd, Galthersburg, MD 20878
            Tel: 3018699056
            Fax: 3018699423
            Email: tdbinfo@tdb.tigr.org
            For clone availability please contact the TIGR Database
            (tdbinfo@tdb.tigr.org)
            Seq primer: M13-21.
            Location/Qualifiers
              1..370
                /organism="Saccharomyces cerevisiae"
                /strain="X2180-1A"
                /db_xref="taxon:4932"
                /clone_1lb="S. cerevisiae strain X2180-1A"
                /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
/

FEATURES
  source
    1..370
      /organism="Saccharomyces cerevisiae"
      /strain="X2180-1A"
      /db_xref="taxon:4932"
      /clone_1lb="S. cerevisiae strain X2180-1A"
      /note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
/

```

BASE COUNT	115 a	73 c	66 g	111 t	5 others
ORIGIN	XhoI"				
Alignment Scores:					
Pred. No.:	2,32e-05	Length:	370		
Score:	124.00	Matches:	21		
Percent Similarity:	90.91%	Conservative:	9		
Best Local Similarity:	63.64%	Mismatches:	3		
Query Match:	27.43%	Indels:	0		
DB:	14	Gaps:	0		
US-09-877-606-4 (1-85) x T36326 (1-370)					
QY	53	LysileysGllleGluasplysllleGluGluileuSerLyslleTyRHsllleGlu	72		
DB	197	AGAAATGAACAACCTTGAGACACAGGTGTGAGAAATTCCTTCGNAATAATATACACTTGCA	138		
QY	73	AsngluilEalAArgllleLysLysleuileGlyGuatg	85		
DB	137	AATGAGTTGCCAGATTAAAGAAATTAAGTTCGCCAAGC	99		
RESULT 3					
T36997/c					
LOCUS	T36997	323 bp	mRNA	linear	EST 11-JAN-1995
DEFINITION	EST102040 S. cerevisiae strain X2180-1A Saccharomyces cerevisiae				
ACCESSION	CDNA 3' end, mRNA sequence.				
VERSION	T36997				
KEYWORDS	T36997.1 GI:620814				
SOURCE	EST				
ORGANISM	baker's yeast.				
	Saccharomyces cerevisiae				
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;				
AUTHORS	Saccharomycetales; Saccharomycetaceae; Saccharomyces.				
TITLE	1 (bases 1 to 323)				
JOURNAL	Weinstock, K.				
COMMENT	Saccharomyces cerevisiae cDNAs				
	Unpublished (1995)				
	Other-ESTs: TC473				
	Contact: Weinstock, K. and Venter, J.C.				
	The Institute for Genomic Research				
	932 Clopper Rd, Galthersburg, MD 20878				
	Tel: 3018699056				
	Fax: 3018699423				
	Email: tdbinfo@tdb.tigr.org				
	For clone availability please contact the TIGR Database				
	(tdbinfo@tdb.tigr.org)				
	Seq primer: M13-21.				
FEATURES					
source	1. 323				
	Location/Qualifiers				
	/organism="Saccharomyces cerevisiae"				
	/strain="X2180-1A"				
	/db_xref="taxon:4932"				
	/clone_lib="S. cerevisiae strain X2180-1A"				
	/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:				
	XhoI"				
BASE COUNT	104 a	57 c	61 g	96 t	5 others
ORIGIN					
Alignment Scores:					
Pred. No.:	3.3e-05	Length:	323		
Score:	122.00	Matches:	21		
Percent Similarity:	90.62%	Conservative:	8		
Best Local Similarity:	65.62%	Mismatches:	3		
Query Match:	26.99%	Indels:	0		
DB:	14	Gaps:	0		
US-09-877-606-4 (1-85) x T36997 (1-323)					
QY	54	IleYssGllleGluasplysllleGluLysleuSerLyslleTyRHsllleGluasn	73		
DB	201	ATGAACAACCTTGAGACACAGGTGTGAGAAATTCCTTCGNAATAATATACACTTGCA	142		

Qy 74 Glu11eAlaAarg11eLysLysLeu11eGlyGluArg 85  
Db 141 GAGGTTCGACATTAAAGAAATTAGTTGGCAACGC 106

RESULT 4  
T36351/c 363 bp mRNA linear EST 11-JAN-1995

LOCUS  
DEFINITION EST101276 S. cerevisiae strain X2180-1A Saccharomyces cerevisiae  
ACCESSION T36351  
VERSION T36351.1 GI:620168  
KEYWORDS  
SOURCE baker's yeast.  
ORGANISM Saccharomyces cerevisiae  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE  
AUTHORS Weinstein, K.  
TITLE Saccharomyces cerevisiae cDNAs  
JOURNAL Unpublished (1995)  
COMMENT Other ESTs: TC473  
Contact: Weinstein, K. and Venter, J.C.  
The Institute for Genomic Research  
932 Clopper Rd, Gaithersburg, MD 20878  
Tel: 3018699056  
Fax: 3018699423  
Email: tdbinfo@tdb.tigr.org  
For clone availability please contact the TIGR Database  
(tdbinfo@tdb.tigr.org)  
Seq primer: M13-21.

FEATURES  
source Location/Qualifiers  
1..363  
/organism="Saccharomyces cerevisiae"  
/strain="X2180-1A"  
/db\_xref="taxon:4932"  
/clone\_lib="S. cerevisiae strain X2180-1A"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"

BASE COUNT 110 a 70 c 71 g 108 t 4 others

ORIGIN

Alignment Scores:  
Pred. No.: 0.000279 Length: 363  
Score: 115.00 Matches: 20  
Percent Similarity: 87.88% Conservative: 9  
Best Local Similarity: 60.61% Mismatches: 4  
Query Match: 25.44% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-4 (1-85) x T36351 (1-363)

Qy 53 Lys11eLysGln11eGluAspLys11eGluLeuSerLys11eTyRHis11eGlu 72  
Db 197 AGATGAAACACTGTGACAGGTTGAAGATTGCTTCGAAATAATATCCTTGAA 138

Qy 73 AsnGlu11eAlaAarg11eLysLysLeu11eGlyGluArg 85  
Db 137 AATGAGTTGCCAGATTAAAGAAATTAGTTGGCAACGC 99

RESULT 5  
T3672/c 210 bp mRNA linear EST 11-JAN-1995

LOCUS  
DEFINITION EST104198 S. cerevisiae strain X2180-1A Saccharomyces cerevisiae  
ACCESSION T3672  
VERSION T3672.1 GI:622489  
KEYWORDS  
SOURCE baker's yeast.  
ORGANISM Saccharomyces cerevisiae  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE  
AUTHORS Weinstein, K.

TITLE  
JOURNAL  
COMMENT Saccharomyces cerevisiae cDNAs  
Unpublished (1995)  
Other ESTs: TC473  
Contact: Weinstein, K. and Venter, J.C.  
The Institute for Genomic Research  
932 Clopper Rd, Gaithersburg, MD 20878  
Tel: 3018699056  
Fax: 3018699423  
Email: tdbinfo@tdb.tigr.org  
For clone availability please contact the TIGR Database  
(tdbinfo@tdb.tigr.org)  
Seq primer: M13-21.

FEATURES  
source Location/Qualifiers  
1..210  
/organism="Saccharomyces cerevisiae"  
/strain="X2180-1A"  
/db\_xref="taxon:4932"  
/clone\_lib="S. cerevisiae strain X2180-1A"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"

BASE COUNT 75 a 36 c 26 g 68 t 5 others

ORIGIN

Alignment Scores:  
Pred. No.: 0.000161 Length: 210  
Score: 114.00 Matches: 19  
Percent Similarity: 87.88% Conservative: 10  
Best Local Similarity: 57.58% Mismatches: 4  
Query Match: 25.22% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-4 (1-85) x T3672 (1-210)

Qy 53 Lys11eLysGln11eGluAspLys11eGluLeuSerLys11eTyRHis11eGlu 72  
Db 192 AGATGAAACACTGTGACAGGTTGAGAAATTGCTTCGAAATAATATCCTTGAA 133

Qy 73 AsnGlu11eAlaAarg11eLysLysLeu11eGlyGluArg 85  
Db 132 AATGAGTTGCCAGATTAAAGAAATTAGTTGGCAACGC 94

RESULT 6  
T37110/c 263 bp mRNA linear EST 11-JAN-1995

LOCUS  
DEFINITION EST102177 S. cerevisiae strain X2180-1A Saccharomyces cerevisiae  
ACCESSION T37110  
VERSION T37110.1 GI:620927  
KEYWORDS  
SOURCE EST.  
ORGANISM Saccharomyces cerevisiae  
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
Saccharomycetales; Saccharomycetaceae; Saccharomyces.

REFERENCE  
AUTHORS Weinstein, K.  
TITLE Saccharomyces cerevisiae cDNAs  
JOURNAL Unpublished (1995)  
COMMENT Other ESTs: TC473  
Contact: Weinstein, K. and Venter, J.C.  
The Institute for Genomic Research  
932 Clopper Rd, Gaithersburg, MD 20878  
Tel: 3018699056  
Fax: 3018699423  
Email: tdbinfo@tdb.tigr.org  
For clone availability please contact the TIGR Database  
(tdbinfo@tdb.tigr.org)  
Seq primer: M13-21.

FEATURES  
source Location/Qualifiers  
1..263  
/organism="Saccharomyces cerevisiae"  
/strain="X2180-1A"  
/db\_xref="taxon:4932"  
/clone\_lib="S. cerevisiae strain X2180-1A"

```

/ncore="Vector: pbluescript SK-, Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      82 a      48 c      43 g      81 t      9 others
ORIGIN

Alignment Scores:
Pred. No.:      0.0016      Length:      263
Percent:      107.00      Matches:      19
Score Similarity:      84.85%      Conservative:      9
Best Local Similarity:      57.58%      Mismatches:      5
Query Match:      13.67%      Indels:      0
DB:      14      Gaps:      0

US-09-877-606-4 (1-85) x T37110 (1-263)

OY      53      LysilleySGNIIleGIIAspLYsIIleGIuGIIleSeuLYsIIeTYrHisIIleGIu 72
Db      197      AGAAGAACAACACTTGAAGACACAGGTTGACAGATTGCTTCGNAATAATTCNCCTTGNA 138
OY      73      ASnguIIleAaArgIIleLYsLYsLeuIIleGIuGIIuArg 85
Db      137      AATGAGTTCACGATTTAAAGAAATTAGTTGGCGACCC 99

RESULT 7
LOCUS      BM166231      747 bp      mRNA      linear      EST 04-DEC-2001
DEFINITION      E57568754 PyBS Plasmodium yoelii yoelii cDNA clone pICNE75 5' end,
ACCESSION      BM166231
VERSION      BM166231.1
KEYWORDS      GI:17299463
SOURCE      EST.
ORGANISM      Plasmodium yoelii yoelii.
AUTHORS      Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
TITLE      1 (bases 1 to 747)
JOURNAL      Carlton, J.M., Dally, T.M., Long, C.A., Bergman, L.W., Valdeya, A.B.,
COMMENT      Fraser, C.M. and Carucci, D.J.
              Plasmodium yoelii EST project at TIGR
              Unpublished (2001)
              Contact: Jane Carlton
              Parasite Genomics Group
              The Institute for Genomic Research
              9712 Medical Center Drive, Rockville, MD 20850, USA
              Tel: 301-530-9319
              Fax: 301-838-0208
              Email: carlton@tigr.org
              For clone info, please contact the Malaria Research and Reference
              Reagent Resource Center, ATCC
              http://www.malaria.mr4.org/mr4pages/index.html
              Seq primer: ADF.

FEATURES
         source
         location/Qualifiers
         1..747
         /organism="Plasmodium yoelii yoelii"
         /strain="17XL"
         /db_xref="taxon:73239"
         /clone="pICNE75"
         /clone_lib="PyBS"
         /dev_stage="asexual blood stages"
         /lab_host="E. coli XL-1 Blue"
         /note="Vector: PAD-GAL4; At 20-25% parasitemia, blood was
         collected from BALB/cByJ mice infected with Py17XL
         parasites, and leukocytes removed by passage over
         microcrystalline cellulose columns. Total RNA was
         isolated using the guanidium/isothiocyanate method, and
         mRNA isolated using oligo(dT)-cellulose chromatography.
         First strand cDNA synthesis was completed using a 50-base
         primer and reverse transcriptase in the presence of
         5-methyl dCTP. After second strand synthesis, uneven
         termini were treated with Pfu DNA polymerase and EcoRI
         adaptors ligated to the blunt ends. The sample was cleaved
         with XhoI and separated on a Sephacryl S-500 column.
         Size-fractionated cDNA was precipitated and ligated to
         HydrizAP arms directionally using EcoRI-XhoI cleaved arms.

```

After packaging, the phagemid vector (PAD-GAL4) was excised from the HybridAmp vector and plasmid DNA isolated."					
BASE COUNT	356 a	85 c	79 g	227 t	
ORIGIN					
<hr/>					
Alignment Scores:					
Pred. No.:	2.77	Length:	747		
Percent Similarity:	86.00	Matches:	26		
Best Local Similarity:	49.33%	Conservative:	11		
Query Match:	34.67%	Mismatches:	26		
DB:	19.03%	Indels:	12		
	13	Gaps:	3		
<hr/>					
US-09-877-606-4 (1-85) x BM166231 (1-747)					
QY	13	AasnAnTYrrThSerLeuIlleHISerLeu-----llegluGluSerGlInAsnnglnGln	30		
Db	46	AAACATGTGGAAAATAATCATTAATCAATTGCAGAACATAGAANAAGATTACAATTCTTCA	105		
QY	31	GIUYSASngluGlnGluLeuLeuLnuksplySTrPhAlaserLeutPrasnrpphe	50		
Db	106	GAGGAATAATATATCATTTCAATCTAAAAAAGCATAAGACTAATAACATTT	165		
QY	51	AsniIeyslgillegllusplSYllegluuilileuSertlySlleTyrrhs	70		
Db	166	AATACTGAAATTAATAAACATTTGAAATTAATAATAGAA-----	204		
QY	71	IlegluAsngluillealargIIelSYslYserIleegluarg	85		
Db	205	--AAAATTAATTTA-----ATTGTAAATTAATATAGAGATAGA	240		
<hr/>					
RESULT 8					
BP027004					
LOCUS					
DEFINITION	BP027004 K. Inaba unpublished cDNA library, testis Clona	284 bp	mRNA	linear	EST 15-MAR-2002
ACCESION	Inestinalis CDNA clone cits46k21 5', mRNA sequence.				
VERSION	BP027004				
KEYWORDS	BP027004.1 GI:19518954				
SOURCE	EST.				
ORGANISM	Ciona intestinalis.				
	Ciona intestinalis.				
	Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;				
	phlebobranchia; Cionidae; Ciona.				
REFERENCE	1 (bases 1 to 284)				
AUTHORS	Inaba,K., Padma,P., Satoh,Y., Shin-i,T., Kohara,Y., Satoh,N. and Satou,Y. EST analysis of gene expression in testis of the ascidian Ciona intestinalis Mol. Reprod. Dev., (2002) In press				
JOURNAL	Contact: Kazuo Inaba				
COMMENT	Asamushi Marine Biological Statio Tohoku University Asamushi, Aomori, Aomori 039-3501, Japan Tel: 81-17-752-3394 Fax: 81-17-752-2765 Email: lnabebiology.tohoku.ac.jp. Location/Qualifiers				
<hr/>					
FEATURES					
source	1..284				
	/organism="Ciona intestinalis"				
	/db_xref="taxon:7719"				
	/clone_cits46k21"				
	/cione_1lb="K. Inaba unpublished cDNA library, testis"				
	/tissue_type="testis"				
	/note="Vector: pZLI: Site 1: EcorI"				
BASE COUNT	122 a	50 c	62 g	49 t	1 others
ORIGIN					
<hr/>					
Alignment Scores:					
Pred. No.:	1.28	Length:	284		
Percent Similarity:	83.50	Matches:	26		
Best Local Similarity:	54.43%	Conservative:	17		
	32.91%	Mismatches:	25		

Query Match: 18.47% Indels: 11  
DB: 14 Gaps: 4

US-09-877-606-4 (1-85) x BP027004 (1-284)

QY 11 GluileAsnAsnTyrThrSerLeuileHisSerLeu-----llegluGlu---Ser 26  
|||:|||||  
DB 39 GAAGTGAATTAACATACACAGCAACTTACTCTGATTCTGTGATGAAGAACACAC 98  
|||:|||||  
QY 27 GluAsnGlnGluLysAsnGlnGluLeuLeuGluLeuAspLysTrpAlaSerLeu 46  
|||:|||||  
DB 99 CAACAGACGACGACATACAGTGAAGATGAAGCAACAGCAAGCAATGCATGAAT 158  
|||:|||||  
QY 47 TrpAspTrpPheAsnLysLysLysGlnLlegluAspLysllegluGluArg 66  
|||:|||||  
DB 159 TCACAC-----CGGATACGACCAATCGCGAAGAAAGCTGAAGAAAGAA 206  
|||:|||||  
QY 67 LysileTyrHisllegluAsnGluLeuAlaArgileLysLysLeuileGlyGluArg 85  
|||:|||||  
DB 207 CAG-----ATAGACACAGCTAGCTGAGTGTGACAAACCAATCAGACGCA 254  
|||:|||||

RESULT 9  
BH280804/c 601 bp DNA linear GSS 30-NOV-2001  
LOCUS BH280804  
DEFINITION CH230-71L16.TV CHORI-230 Segment 1 Rattus norvegicus genomic clone  
CH230-71L16, DNA sequence.  
ACCESSION BH280804  
VERSION BH280804.1 GI:17193206  
KEYWORDS GSS.  
SOURCE Norway rat.  
ORGANISM Rattus norvegicus  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 601)  
Zhao,S., Shetty,D., Shatsman,S., Tsegaye,G., Geer,K., Shvartsbeyn  
A., Gabregorgis,E., Overton,L., Russell,D., Chen,D., Rllygs,F., de  
Jong,P. and Fraser,C.M.  
Rat BAC End Sequences from Library CHORI-230 EcoRI segment  
Unpublished (1999) CH230-71L16.TJ  
Other\_GSSs: CH230-71L16.TJ  
Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rat230.htm). For BAC library  
availability, please contact Pieter de Jong (pjejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/or ering\_informtion.htm). BAC end  
page: http://www.tigr.org/tdb/bac-ends/rat/bac\_end\_intro.html  
Plate: 71 row: L column: 16  
Seq primer: T7  
Class: BAC ends.

FEATURES  
Source Location/Qualifiers  
1..601  
/organism="Rattus norvegicus"  
/strain="BN/SSNHSd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-71L16"  
/clone\_1ib="CHORI-230 Segment 1"  
/sex="female"  
/cell\_type="Brain"  
/note="Vector: pTARBAC2.1; Site\_1: EcoRI; Site\_2: EcoRI;  
CHORI-230 Rat (BN/SSNHSd/MCW) BAC library produced by  
Pieter de Jong"

BASE COUNT 118 a 106 c 105 g 272 t

Alignment Scores:

Pred. No.: 4 Length: 601  
Score: 83.50 Matches: 23  
Percent Similarity: 52.388 Conservative: 21  
Best Local Similarity: 27.388 Mismatches: 29  
Query Match: 18.47% Indels: 11  
DB: 17 Gaps: 4

US-09-877-606-4 (1-85) x BH280804 (1-601)

QY 8 TrpAspArgGluileAsnAsnTyrThr---SerLeuileHisSerLeuileGluLys 26  
|||:|||||  
DB 459 TGCGACACACACTAAACACTTATACAAAACCAATACATGCAAGATGAGAAAGC 400  
|||:|||||  
QY 27 Gln-----AsnGlnGlnLysAsnGlnGlnGlnGlnGlnGlnGln 40  
|||:|||||  
DB 399 ATGAATTAACCTCAGGAGTGAAGTCAACCAAGTGAAGTCAAGTGAAGTAC----- 346  
|||:|||||  
QY 41 AspLysTrpAlaSerLeuTrpAsnTrpPheAsnLysLysLysGlnLlegluAspLys 60  
|||:|||||  
DB 345 ---AAGAAATCAACATCTGCAACTGTCTTCTTAAATAATCATGAGATA---GATAAA 292  
|||:|||||  
QY 61 llegluGluileLysSerLysLysLysllegluAsnGluileAlaArgileLysLys 80  
|||:|||||  
DB 291 CTCATAGCCAGACTAACCAGAGGACAGAGTGTATTCAATTAACAAGTCAGAAAT 232  
|||:|||||  
QY 81 LeuileGlyGlu 84  
|||:|||||  
DB 231 GGAGAGGGAGAC 220  
|||:|||||

RESULT 10  
AQ376798/c 541 bp DNA linear GSS 20-MAY-1999  
LOCUS AQ376798/c  
DEFINITION RPCI11-151G19.TV RPCI-11 Homo sapiens genomic clone RPCI-11-151G19,  
DNA sequence.  
ACCESSION AQ376798  
VERSION AQ376798.1 GI:4347821  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
1 (bases 1 to 541)  
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter  
J.C.  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
Unpublished (1997) RPCI11-151G19.TJ  
Other\_GSSs: RPCI11-151G19.TJ  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbeet@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pjejong@med.bufo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.bufo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/tdb/hungen/bac\_end\_search.html  
Seq primer: T7  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..541  
/organism="Homo sapiens"  
/db\_xref="GDB:7557762"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-151G19"  
/clone\_1ib="RPCI-11"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;

BASE COUNT	113 a	96 c	76 g	251 t	5 others
ORIGIN					
Alignment Scores:					
Pred. No.:	3.92	Length:	541		
Score:	83.00	Matches:	27		
Percent Similarity:	47.92%	Conservative:	19		
Best Local Similarity:	28.12%	Mismatches:	20		
Query Match:	18.36%	Indels:	30		
DB:	17	Gaps:	5		
US-09-877-606-4 (1-85) x AQ376798 (1-541)					
OY	14	AsnTyrThrSerLeuIleHisSerLeuIle-----GlucIuSerGlnAsnGln	29		
DB	473	GATTTCNMAAGCCTGATTATTAGTTACTCAGAGAATATACAGAGAAAGTAAACCAA	414		
OY	29	-----	29		
DB	413	CTTANAGAAATTTTAAACATTCAGCATATGAAATTTTCTGCACAGATATAT	354		
OY	30	---GingIuLyAsnGlnGlnIuLeuIeuGlnIeu---AspLysTrpAlaSerIeuTrp	47		
DB	353	ATTTTAAAGAAACCAACACAGACTTCTGGAATATGAAGACATTTAGCGGATTACA	294		
OY	48	Asn-----TrpHeAsnIleLysIleLysGlnIleGluAspLysIleGluIle	64		
DB	293	AATGCAGAGGAAGATTATACAAATAGGTAGCAGCAAGTAGAAGAAAGAAATTTTCAGAGGTC	234		
OY	65	LeuSerLysIleTyrHisIleGluAsnGlnIuLeuAlaArgIleLysLys	80		
DB	233	AAAGCAGAGACTTTGCAATTTACCCACACAG---ACAAATAATAAAGAAA	189		
RESULT 11	AQ793621	543 bp	DNA	linear	GSS 03-AUG-1999
LOCUS	HS_5421_B1_B06_T7A	RPCT-11	Human Male	BAC	Library Homo sapiens
DEFINITION	genomic clone Plate-997 Col-11 Row-D, DNA sequence.				
ACCESSION	AQ793621				
VERSION	AQ793621.1	GI:5701245			
KEYWORDS	GSS.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
	1 (bases 1 to 543)				
	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,				
	Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and				
	Hood,L.				
	Sequence-tagged connectors: A sequence approach to mapping and				
	scanning the human genome				
	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)				
JOURNAL	99380589				
MEDLINE	Contact: Mahairas GG, Wallace JC, Hood L				
COMMENT	High Throughput Sequencing Center				
	University of Washington				
	401 Queen Anne Avenue North, Seattle, WA 98109, USA				
	Tel: (206) 616-3618				
	Fax: (206) 616-3887				
	Email: jwallace@u.washington.edu				
	Clones are derived from the human BAC library RPCT-11. For BAC				
	library availability, please contact Pieter de Jong				
	(pieter@dejong.med.buffalo.edu). Clones may be purchased from				
	BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)				
	or from Resear h Genetics (info@resgen.com). BAC end Web Server:				
	http://www.hsc.washington.edu				
	plate: 997 row: D column: 11				
	Seq primer: 17				
	Class: BAC ends				
	High quality sequence stop: 543.				
FEATURES	location/Qualifiers				
Source	1..543				

			/organism="Homo sapiens"
		/db.xref="taxon:9606"	
		/clone="Plate-997 Col-11 Row-D"	
		/clone_1lb="RPCT-11 Human Male BAC Library"	
		/sex="male"	
		/note="Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at EcoRI sites"	
BASE COUNT	233 a	73 c	111 g 122 t 4 others
ORIGIN			
Alignment Scores:			
Pred. No.:	3.95	Length:	543
Score:	83.00	Matches:	25
Percent Similarity:	53.66%	Conservative:	19
Best Local Similarity:	30.49%	Mismatches:	30
Query Match:	18.36%	Indels:	8
DB:	17	Gaps:	3
US-09-877-606-4 (1-85) x AQ793621 (1-543)			
Oy	8	TTPAARGGLIIEASNAstYrThrserLeulleHSerLeulleGIUGLusergin	27
Dd	37	TGGGAGCtCAAAATTCAGATCACAGAAAAGATCATATTCTTGCAGATTAAATTAAT	96
Oy	28	AsngInglIngLu-----LysAnsglIngLuLeuLeuGluleu-----	40
Dd	97	GAAGAGATTGAATAATTTAAAGAATGAATAAGAAATTCGTGAGTTGAAAAATGCCAATT	156
Oy	41	AsplySTPAlaselerLeutrpasnTrpheanIllelySIllelyglnlleluaspLyS	60
Dd	157	GACACATGAAGAATGATCGGAGTCCTTCCAACAGCAGCATTGATGATCAACAAGAAAA	216
Oy	61	IlegluGuilleSerlySlleTy---HSileguAnsnguIlleAlarglleyS	79
Dd	217	ATTAGTAGCTCAAGACAGGATATTTCGAAATACACACTGCAGAGAGCAAGAAAA	276
Oy	80	LysLeu	81
Dd	277	AGATA	282
RESULT 12			
LOCUS	A1676261		
DEFINITION	A1676261 Etch1 Eimeria tenella CDNA clone etma003 5', mRNA		
ACCSSION	A1676261		
VERSION	A1676261.1 GI:4876741		
KEYWORDS	EST.		
SOURCE	Eimeria tenella.		
ORGANISM	Eimeria tenella.		
REFERENCE	Eukaryota: Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;		
AUTHORS	Eimeria.		
TITLE	1 (bases 1 to 406)		
JOURNAL	Wan,K.L., Chong,S.P., Ng,S.T., Tomley,F.M. and Jangl,M.S.		
COMMENT	Survey of genes in Eimeria tenella metozoite		
	Unpublished (1999)		
	Contact: Wan KL		
	Centre for Gene Analysis and Technology		
	Universiti Kebangsaan Malaysia		
	43600 UKM Bangi, Selangor DE, Malaysia		
	Tel: 6 03 8292997		
	Fax: 6 03 8293249		
	Email: klian@pkrtisc.cc.ukm.my		
	PCR Primers		
	FORWARD: T3		
	BACKWARD: T7		
	Sed primer: SK.		
FEATURES	Location/Qualifiers		
SOURCE	1..406		
	/organism="Eimeria tenella"		

```

/dey stage="asexual blood stages"
/lab host="E. coli XL-1 Blue"
/note="Vector: PAD-GAL4; At 20-25% parasitemia, blood was collected from BALB/cByJ mice infected with Pyl17XL parasites, and leukocytes removed by passage over microcrystalline cellulose columns. Total RNA was isolated using the guanidium isothiocyanate method, and mRNA isolated using oligo(dT)-cellulose chromatography. First strand cDNA synthesis was completed using a 50-base primer and reverse transcriptase in the presence of 5-methyl dCTP. After second strand synthesis, uneven termini were treated with Pfu DNA polymerase and EcoRI adaptors ligated to the blunt ends. The sample was cleaved with XhoI and separated on a Sephacryl S-500 column. Size-fractionated cDNA was precipitated and ligated to Hydrizap arms directionally using EcoRI-XhoI cleaved arms. After packaging, the phagemid vector (PAD-GAL4) was excised from the Hydrizap vector and plasmid DNA isolated."

BASE COUNT      177 a       29 c       92 g       46 t
ORIGIN

Alignment Scores:
Pred. NO.:          2.61           Length:         344
Score:             82.00           Matches:        16
Percent Similarity: 58.73%         Conservative:   21
Best local Similarity: 25.40%       Mismatches:    16
Query Match:       18.14%          Indels:         10
DB:                13              Gaps:           1

US-09-877-606-4 (1-85) x BM170439 (1-344)

QY 23 IlegGluSerGlnAsnGInGInGluLysAsnGInGInGluLeuLeuGluDeuAspLys 42
:: ||| ::|||:: :::: ||:::|||| |::
Db 65 GTAACCGAGAGCAAAATAAAGTGAAAGAGAAGAAAAAATGTGAAGAGAGAGAATAAT 124

QY 43 TrpAlaSerLeuTPAsnTrpPheasnIleLysIleTyIeLysGInIleGlaAspLysIleGu 62
||| ::|||:: :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 125 -----AAAGTGAAGAGATCGAAAACAAGTTGAA 154
||| ::|||:: :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
QY 63 GluIleLeuSerIlyIleTyrlHisIleGluAsnGluIleAlaIleArgIleLysLysLeuIle 82
||| ::|||:: :|||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
Db 155 GAGAGAGAAAACAGACGTACAGACATGCAGAAAACAAGTAAGTAAGACATGTGAAGCCAAAGTG 214

QY 83 GlytIuarG 85
|||:::
Db 215 GCCGAAAAA 223

RESULT 14
LOCUS      BM170211               364 bp     mRNA          linear      EST 04-DEC-2001
DEFINITION ESr572734 PYBS Plasmodium yoelii yoelii cDNA clone pYCQC93 5' end,
ACCESSION  BM170211
KEYWORDS   mRNA sequence.
SOURCE     BM170211.1 GI:17303443
ORGANISM   EST.
            Plasmodium yoelii yoelii.
            Plasmodium yoelii yoelii.
            Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
AUTHORS    1 (bases 1 to 364)
            Carlton,J.M., Daly,T.M., Long,C.A., Bergman,L.W., Valdyg,A.B.,
            Fraser,C.M. and Carucci,D.J.
TITLE      Plasmodium yoelii EST project at TIGR
COMMENT    Unpublished (2001)
            Contact: Jane Carlton
            Parasite Genomics Group
            The Institute for Genomic Research
            9712 Medical Center Drive, Rockville, MD 20850, USA
            Tel: 301-530-9319
            Fax: 301-838-0208
            Email: carltonetigr.org
            For clone info, please contact the Malaria Research and Reference
            Reagent Resource Center, ATCC

```





GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:27:06 ; Search time 5 Seconds  
(without alignments)  
186.551 Million cell updates/sec

Title: US-09-877-606-7

Perfect score: 48

Sequence: 1 LMWFMNI 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: A\_Geneseq.101002.\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the change being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	48	100.0	15	10 AAP90238	Antigenic peptide
2	48	100.0	15	22 AAU12524	Human HIV-1 Th-CTL
3	48	100.0	15	22 AAU12532	Human HIV-1 Th-CTL
4	48	100.0	15	22 AAU12542	Human HIV-1 Th-CTL
5	48	100.0	20	17 AAU07920	gp41 peptide 53.
6	48	100.0	23	9 AAP82469	Peptide component
7	48	100.0	23	17 AAU07974	HIV peptide 41-21.
8	48	100.0	23	22 AAU12483	HIV Th-CTL peptide
9	48	100.0	24	22 AAU12488	HIV Th-CTL p17 epi
10	48	100.0	24	23 AAU070251	HIV Th-CTL p17 epi

11	48	100.0	32	22 AAU12479	HIV Th-CTL peptide
12	48	100.0	32	23 AAU70242	HIV Th-CTL overlap
13	48	100.0	33	23 ABB09199	HIV-1 gp160 amino
14	48	100.0	36	14 AAU17022	DP-178-derived pep
15	48	100.0	36	21 AAU89188	Core polypeptide f
16	48	100.0	36	21 AAU89189	Core polypeptide f
17	48	100.0	36	21 AAU89190	Core polypeptide f
18	48	100.0	36	21 AAU89191	Core polypeptide f
19	48	100.0	36	21 AAU89193	Core polypeptide f
20	48	100.0	36	21 AAU89194	Core polypeptide f
21	48	100.0	36	21 AAU89195	Core polypeptide f
22	48	100.0	36	21 AAU89196	Core polypeptide f
23	48	100.0	36	21 AAU89197	Core polypeptide f
24	48	100.0	36	21 AAU89198	Core polypeptide f
25	48	100.0	36	21 AAU89199	Core polypeptide f
26	48	100.0	36	21 AAU89200	Core polypeptide f
27	48	100.0	36	22 ABB00546	Core polypeptide f
28	48	100.0	36	22 ABB00547	Core polypeptide f
29	48	100.0	36	22 ABB00548	Core polypeptide f
30	48	100.0	36	22 ABB00549	Core polypeptide f
31	48	100.0	36	22 ABB00550	Core polypeptide f
32	48	100.0	36	22 ABB00551	Core polypeptide f
33	48	100.0	36	22 ABB00552	Core polypeptide f
34	48	100.0	36	22 ABB00553	Core polypeptide f
35	48	100.0	36	22 ABB00554	Core polypeptide f
36	48	100.0	36	22 ABB00555	Core polypeptide f
37	48	100.0	36	22 ABB00556	Core polypeptide f
38	48	100.0	36	22 ABB00557	Core polypeptide f
39	48	100.0	36	22 ABB00558	Core polypeptide f
40	48	100.0	36	22 ABB02022	Core polypeptide f
41	48	100.0	36	22 ABB02023	Core polypeptide f
42	48	100.0	36	22 ABB02024	Core polypeptide f
43	48	100.0	36	22 ABB02025	Core polypeptide f
44	48	100.0	36	22 ABB02026	Core polypeptide f
45	48	100.0	36	22 ABB02027	Core polypeptide f

## ALIGNMENTS

RESULT 1  
ID AAP90238 standard; peptide; 15 AA.  
XX  
AC AAP90238:  
XX  
DT 26-FEB-1990 (first entry)  
XX  
DE Antigenic peptide for detecting, inhibiting and neutralising HIV-1.  
XX  
KW HIV-1; antigenic determinants.  
XX  
OS HIV-1.  
XX  
PN EP330359-A.  
XX  
PD 30-AUG-1989.  
XX  
PF 14-FEB-1989; 89EP-0301364.  
XX  
PR 25-FEB-1988; 88US-0160378.  
XX  
PA (BIRA ) BIO RAD LABS INC.  
XX  
PI Walker RP, Parekh BS;  
XX  
DR WPI; 1989-250452/35.  
XX  
PT New peptide(s) for detecting, inhibiting and neutralising HIV-1  
PT corresp. to antigenic determinants encoded by conserved regions  
PT of HIV-1 genome.  
XX  
PS Disclosure; page 4; 16pp; English.

XX Peptide is one of several fragments from gp 120 and gp41 (env gene)  
CC p18 and p24 (gag gene), p32 (pol gene) and proteins encoded by the  
CC tat, rev, trs/art and sor genes. They are used for detecting,  
CC inhibiting and neutralising HIV-1 infection. Dosage is pref.  
CC 5-25 mg/kg. The peptides can be used for any type of immunological  
CC detection esp. dot blot and ELISA.  
CC See also AAP90191-P90274.  
XX  
SQ Sequence 15 AA:  
Query Match 100.0%; Score 48; DB 10; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LMNMFNI 7  
| | | | |  
DB 9 LMNMFNI 15  
RESULT 2  
AAU12524  
ID AAU12524 standard; Peptide: 15 AA.  
XX  
AC AAU12524;  
XX  
DT 27-SEP-2001 (first entry)  
XX  
DE Human HIV-1 Th-CTL overlapping epitope #31.  
XX  
KW Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;  
KW human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;  
KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;  
KW Vaccinia ankara.  
XX  
OS Homo sapiens.  
OS Human immunodeficiency virus type 1.  
XX  
FN WO200156355-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US03540.  
XX  
PR 04-FEB-2000; 2000US-0497497.  
XX  
PA (UYDU-) UNIV DUKE.  
XX  
PI Haynes BF, Liao H;  
XX  
DR WPI: 2001-488827/53.  
XX  
PT Vaccine for immunizing against human immunodeficiency virus has mixture  
PT or linear array of peptides comprising immunodominant T-helper epitopes  
PT and major histocompatibility complex cytotoxic T-lymphocyte epitopes -  
XX  
PS Claim 1; Page 26; 33pp; English.  
XX  
CC The present invention relates to human immunodeficiency virus (HIV)  
CC and in particular to a human leukocyte antigen (HLA)-based HIV vaccine.  
CC The vaccine comprises a mixture or linear array of peptides, or its  
CC variants, where the peptides contain immunodominant T-helper (Th)  
CC epitopes and major histocompatibility complex (MHC) cytotoxic T-  
CC lymphocyte (CTL) epitopes and the linear array of peptides are  
CC preferably expressed in modified Vaccinia ankara. The vaccine is useful  
CC for immunising a patient against HIV and focuses immune response on many  
CC dominant and subdominant CTL epitopes of HIV. DNA or live vectors with  
CC linear arrays of CTL epitopes can be used as either primes or boosts of  
CC peptides or of each other to optimally give CTL anti-HIV responses. The  
CC vaccine induces salutary anti-HIV immune responses. AAU12447-AAU12558  
CC represent the amino acid sequences of the Th-CTL epitopes and HIV  
XX immunogenic peptides used in the invention.

SQ Sequence 15 AA:  
Query Match 100.0%; Score 48; DB 22; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LMNMFNI 7  
| | | | |  
DB 3 LMNMFNI 9  
RESULT 3  
AAU12532  
ID AAU12532 standard; Peptide: 15 AA.  
XX  
AC AAU12532;  
XX  
DT 27-SEP-2001 (first entry)  
XX  
DE Human HIV-1 Th-CTL overlapping epitope #39.  
XX  
KW Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;  
KW human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;  
KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;  
KW Vaccinia ankara.  
XX  
OS Homo sapiens.  
OS Human immunodeficiency virus type 1.  
XX  
FN WO200156355-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 05-FEB-2001; 2001WO-US03540.  
XX  
PR 04-FEB-2000; 2000US-0497497.  
XX  
PA (UYDU-) UNIV DUKE.  
XX  
PI Haynes BF, Liao H;  
XX  
DR WPI: 2001-488827/53.  
XX  
PT Vaccine for immunizing against human immunodeficiency virus has mixture  
PT or linear array of peptides comprising immunodominant T-helper epitopes  
PT and major histocompatibility complex cytotoxic T-lymphocyte epitopes -  
XX  
PS Disclosure; Page 27; 33pp; English.  
XX  
CC The present invention relates to human immunodeficiency virus (HIV)  
CC and in particular to a human leukocyte antigen (HLA)-based HIV vaccine.  
CC The vaccine comprises a mixture or linear array of peptides, or its  
CC variants, where the peptides contain immunodominant T-helper (Th)  
CC epitopes and major histocompatibility complex (MHC) cytotoxic T-  
CC lymphocyte (CTL) epitopes and the linear array of peptides are  
CC preferably expressed in modified Vaccinia ankara. The vaccine is useful  
CC for immunising a patient against HIV and focuses immune response on many  
CC dominant and subdominant CTL epitopes of HIV. DNA or live vectors with  
CC linear arrays of CTL epitopes can be used as either primes or boosts of  
CC peptides or of each other to optimally give CTL anti-HIV responses. The  
CC vaccine induces salutary anti-HIV immune responses. AAU12447-AAU12558  
CC represent the amino acid sequences of the Th-CTL epitopes and HIV  
XX immunogenic peptides used in the invention.  
SQ Sequence 15 AA:  
Query Match 100.0%; Score 48; DB 22; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LMNMFNI 7  
| | | | |  
DB 3 LMNMFNI 9

```

RESULT 4
AAU12542
ID AAU12542 standard; Peptide; 15 AA.
XX
AC AAU12542;
XX
DT 27-SEP-2001 (first entry)
XX
DE Human HIV-1 Th-CTL overlapping epitope #49.
XX
KW Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW Vaccinia ankara.
XX
OS Homo sapiens.
OS Human immunodeficiency virus type 1.
XX
PN MO200156355-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03540.
XX
PR 04-FEB-2000; 2000US-0497497.
XX
PA (UYDU-) UNIV DUKE.
XX
PI Haynes BF, Liao H;
XX
DR WPI; 2001-488827/53.
XX
PT Vaccine for immunizing against human immunodeficiency virus has mixture
PT or linear array of peptides comprising immunodominant T-helper epitopes
XX and major histocompatibility complex cytotoxic T-lymphocyte epitopes
XX
PS Disclosure; Page 27; 33pp; English.
XX
CC The present invention relates to human immunodeficiency virus (HIV)
CC and in particular to a human leukocyte antigen (HLA)-based HIV vaccine.
CC The vaccine comprises a mixture or linear array of peptides, or its
CC variants, where the peptides contain immunodominant T-helper (Th)
CC lymphocyte (CTL) epitopes and the linear array of peptides are
CC preferably expressed in modified Vaccinia ankara. The vaccine is useful
CC for immunising a patient against HIV and focuses immune response on many
CC dominant and subdominant CTL epitopes of HIV. DNA or live vectors with
CC linear arrays of CTL epitopes can be used as either primes or boosts of
CC peptides or of each other to optimally give CTL anti-HIV responses. The
CC vaccine induces salutary anti-HIV immune responses. AAU12447-AAU12558
CC represent the amino acid sequences of the Th-CTL epitopes and HIV
CC immunogenic peptides used in the invention.
XX
SQ Sequence 15 AA;
Query Match 100.0%; Score 48; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LMNMFNI 7
| | | | |
DB 3 LMNMFNI 9

```

```

XX
DE gp41 peptide 53.
XX
KW HIV; gp120; HIV-IIIB strain; HIV-1 transmission; foetal transmission;
KW neutralising antibody; passive immunisation; anti-idiotypic antibody;
KW gp41; vaccine; active immunotherapy.
XX
OS Human immunodeficiency virus type 1.
XX
PN US5556744-A.
XX
PD 17-SEP-1996.
XX
PF 24-MAR-1994; 94US-0218025.
XX
PR 24-MAR-1994; 94US-0218025.
PR 29-MAY-1992; 92US-0891451.
XX
PA (UYPE-) UNIV PENNSYLVANIA.
PA (WIST-) WISTAR INST ANATOMY & BIOLOGY.
XX
PI Ugen KE, Weiner DB, Williams WV;
XX
DR WPI; 1996-432980/43.
XX
PT Determining the likelihood of maternal transmission of HIV-1 to
PT foetus - by measuring maternal reactivity with specific gp120 and
PT gp41 derived peptide(s), also used for diagnosing HIV in infants
XX
PS Claim 1; Column 115; 63pp; English.
XX
CC AAU07918-W07928 represent HIV gp41 peptides that can be used in the
CC method of the invention. These sequences are derived from the HIV-IIIB
CC strain deposited as ATCC CRL 8543. The method of the invention is for
CC determining whether or not a mother will transmit HIV-1 to a foetus.
CC The method comprises incubating a sample from the HIV-infected mother,
CC with a collection of HIV peptides. The HIV peptides includes at least
CC one of these sequences, and at least one HIV gp120 derived peptide (see
CC AAU07909-W07917). The number of peptides that react with the sample is
CC determined, and this number is compared with a standard that shows
CC pattern reactivity for a patient of transmission status. A
CC non-transmissible HIV sample is indicated if the test sample reacts with
CC twice as many peptides as the standard. The method detects the presence
CC of neutralising antibodies that protect against mother to infant
CC transmission of HIV. These sequences can also be used in vaccines to
CC protect against transmission. Antibodies against these sequences can be
CC used for passive immunisation, and to generate anti-idiotypic antibodies
CC for use in vaccines or active immunotherapy.
XX
SQ Sequence 20 AA;
Query Match 100.0%; Score 48; DB 17; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.54;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LMNMFNI 7
| | | | |
DB 8 LMNMFNI 14

```

```

RESULT 5
AAU07920
ID AAU07920 standard; peptide; 20 AA.
XX
AC AAU07920;
XX
DT 31-JAN-1997 (first entry)

```

```

RESULT 6
AAP82469
ID AAP82469 standard; protein; 23 AA.
XX
AC AAP82469;
XX
DT 12-NOV-1990 (first entry)
XX
DE Peptide component of AIDS vaccine.
XX
KW AIDS vaccine; T-cells.
XX
OS synthetic.

```

XX EP273716-A.  
 PN 06-JUL-1988.  
 PS  
 XX 23-DEC-1987; 87EP-0311391.  
 PF  
 XX 12-FEB-1987; 87US-0014430.  
 PR 30-DEC-1986; 86US-0947935.  
 XX  
 PA (USDC ) US SEC OF COMMERCE.  
 PA (USSH ) US DEPT HEALTH AND HUMAN.  
 PI Delisi C, Margalit H, Cornette JL, Ouyang CS;  
 DR WPI: 1988-184640/27.  
 XX  
 PT Synthetic peptide(s) as vaccines for AIDS - selected from peptide  
 PT regions which can fold as a maximally amphipathic helix recognised  
 PT by T cells.  
 XX  
 PS Claim 9; Page 10; 16pp; English.  
 CC This peptide is a component of an AIDS vaccine. It can fold as a  
 CC maximally amphipathic helix and is recognised by T-cells immune to  
 CC the AIDS virus envelope protein.  
 CC See also AAP82462-68 and AAP82470-79.  
 CC  
 XX Sequence 23 AA;

Query Match 100.0%; Score 48; DB 9; Length 23;  
 Best Local Similarity 100.0%; Pred. NO. 0.62;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMENI 7  
 |||||  
 DB 11 LMNMENI 17

RESULT 7  
 AAM07974  
 ID AAM07974 standard; peptide: 23 AA.  
 XX

AC AAM07974;

DT 03-FEB-1997 (first entry)

XX HIV peptide 41-21.

XX HIV; gp120; HIV-1TB strain; HIV-1 transmission; foetal transmission;  
 KW neutralising antibody; passive immunisation; anti-idiotypic antibody;  
 KW gp41; vaccine; active immunotherapy.  
 XX

OS Human immunodeficiency virus type 1.

XX USS556744-A.

PN 17-SEP-1996.

XX 24-MAR-1994; 94US-0218025.

XX 24-MAR-1994; 94US-0218025.

XX 29-MAY-1992; 92US-0891451.

XX (UYPE-) UNIV PENNSYLVANIA.

XX (WIST-) WISTAR INST ANATOMY & BIOLOGY.

XX Ugen KE, Weiner DB, Williams WV;

XX WPI: 1996-432980/43.

PT Determining the likelihood of maternal transmission of HIV-1 to  
 PT foetus - by measuring maternal reactivity with specific gp120 and

PT gp41 derived peptide(s), also used for diagnosing HIV in infants  
 XX  
 XX Example 6; Column 95-96; 63pp; English.

CC AAM07956-W07979 represent HIV peptides that can be used in the method  
 CC of the invention. The method of the invention is for determining  
 CC whether or not a mother will transmit HIV-1 to a foetus. The method  
 CC comprises incubating a sample from the HIV-infected mother, with a  
 CC collection of HIV peptides. The HIV peptides includes at least one HIV  
 CC gp120 derived peptide (see AAM07909-W07917), and at least one HIV gp41  
 CC derived peptide (see AAM07918-W07928). The number of peptides that react  
 CC with the sample is determined, and this number is compared with a  
 CC standard that shows pattern reactivity for a patient of transmission  
 CC status. A non-transmissible HIV sample is indicated if the test sample  
 CC reacts with twice as many peptides as the standard. The method detects  
 CC the presence of neutralising antibodies that protect against mother to  
 CC infant transmission of HIV. These sequences can also be used in vaccines  
 CC to protect against transmission. Antibodies against these sequences can  
 CC be used for passive immunisation, and to generate anti-idiotypic  
 CC antibodies for use in vaccines or active immunotherapy.  
 CC  
 XX Sequence 23 AA;

Query Match 100.0%; Score 48; DB 17; Length 23;  
 Best Local Similarity 100.0%; Pred. NO. 0.62;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMENI 7  
 |||||  
 DB 9 LMNMENI 15

RESULT 8  
 AAU12483  
 ID AAU12483 standard; Peptide: 23 AA.  
 XX

AC AAU12483;

DT 27-SEP-2001 (first entry)

XX HIV Th-CTL peptide prototype vaccine immunogen #20.

XX Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;  
 KW human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;  
 KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;  
 KW Vaccinia ankara.  
 XX

OS Homo sapiens.

PN W0200156355-A2.

XX 09-AUG-2001.

XX 05-FEB-2001; 2001WO-US03540.

XX 04-FEB-2000; 2000US-0497497.

XX (UYDU-) UNIV DUKE.

XX Haynes BF, Liao H;

XX WPI: 2001-488827/53.

PT Vaccine for immunizing against human immunodeficiency virus has mixture  
 PT or linear array of peptides comprising immunodominant T-helper epitopes  
 PT and major histocompatibility complex cytotoxic T-lymphocyte epitopes -  
 XX  
 XX Claim 1; Page 24; 33pp; English.

CC The present invention relates to human immunodeficiency virus (HIV)  
 CC and in particular to a human leukocyte antigen (HLA)-based HIV vaccine.  
 CC The vaccine comprises a mixture or linear array of peptides, or its  
 CC variants, where the peptides contain immunodominant T-helper (Th)

CC epitopes and major histocompatibility complex (MHC) cytotoxic T-  
 CC lymphocyte (CTL) epitopes and the linear array of peptides are  
 CC preferably expressed in modified Vaccinia ankara. The vaccine is useful  
 CC for immunising a patient against HIV and focuses immune response on many  
 CC dominant and subdominant CTL epitopes of HIV. DNA or live vectors with  
 CC linear arrays of CTL epitopes can be used as either primes or boosts of  
 CC peptides or of each other to optimally give CTL anti-HIV responses. The  
 CC vaccine induces salutary anti-HIV immune responses. AAU12447-AAU12558  
 CC represent the amino acid sequences of the Th-CTL epitopes and HIV  
 CC immunogenic peptides used in the invention.

XX Sequence 23 AA;

Query Match 100.0%; Score 48; DB 22; Length 23;

Best Local Similarity 100.0%; Pred. No. 0.62; Mismatches 0; Gaps 0;

Matches 7; Conservative 0; Indels 0; Gaps 0;

RESULT 9  
 AAU12488  
 ID AAU12488 standard; Peptide; 24 AA.

XX AAU12488;

XX 27-SEP-2001 (first entry)

DE HIV Th-CTL peptide prototype vaccine immunogen #25.

XX Human immunodeficiency virus: HIV; immunogen; vaccine; anti-HIV;  
 KW human leukocyte antigen; HLA: T-helper epitope; CTL; Th; epitope;  
 KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;  
 KW Vaccinia ankara.

XX Homo sapiens.

PN WO200156355-A2.

PD 09-AUG-2001.

PF 05-FEB-2001; 2001WO-US03540.

PR 04-FEB-2000; 2000US-0497497.

PA (UYDU-) UNTV DUKE.

PI Haynes BF, Liao H;

DR WPI; 2001-488827/53.

XX Vaccine for immunizing against human immunodeficiency virus has mixture  
 PT or linear array of peptides comprising immunodominant T-helper epitopes  
 PR and major histocompatibility complex cytotoxic T-lymphocyte epitopes -  
 XX Claim 1; Page 25; 33pp; English.

XX The present invention relates to human immunodeficiency virus (HIV)  
 CC and in particular to a human leukocyte antigen (HLA)-based HIV vaccine.  
 CC The vaccine comprises a mixture or linear array of peptides, or its  
 CC variants, where the peptides contain immunodominant T-helper (Th)  
 CC epitopes and major histocompatibility complex (MHC) cytotoxic T-  
 CC lymphocyte (CTL) epitopes and the linear array of peptides are  
 CC preferably expressed in modified Vaccinia ankara. The vaccine is useful  
 CC for immunising a patient against HIV and focuses immune response on many  
 CC dominant and subdominant CTL epitopes of HIV. DNA or live vectors with  
 CC linear arrays of CTL epitopes can be used as either primes or boosts of  
 CC peptides or of each other to optimally give CTL anti-HIV responses. The  
 CC vaccine induces salutary anti-HIV immune responses. AAU12447-AAU12558  
 CC represent the amino acid sequences of the Th-CTL epitopes and HIV  
 CC immunogenic peptides used in the invention.

XX Sequence 24 AA;

Query Match 100.0%; Score 48; DB 22; Length 24;

Best Local Similarity 100.0%; Pred. No. 0.64; Mismatches 0; Gaps 0;

Matches 7; Conservative 0; Indels 0; Gaps 0;

RESULT 10  
 AAU70251  
 ID AAU70251 standard; Peptide; 24 AA.

XX AAU70251;

XX 14-FEB-2002 (first entry)

DE HIV Th-CTL P17 epitopes (A2 variants) prototype vaccine D-Th/L-CTL.

XX HIV; human immunodeficiency virus; AIDS; SIV; anti-HIV; virucide;  
 KW acquired immunodeficiency syndrome; simian immunodeficiency virus;  
 KW vaccine; cytotoxic T cell response inducer; Th; human leukocyte antigen;  
 KW immunodominant T-helper epitope; cytotoxic T lymphocyte; CTL; HLA.

XX Human immunodeficiency virus type 1.  
 OS Synthetic.

PN US2001036461-A1.

PD 01-NOV-2001.

PF 05-FEB-2001; 2001US-0775805.

PR 04-FEB-2000; 2000US-0497497.

PA (HAYN/) HAYNES B F.

PI (LIAO/) LIAO H.

DR WPI; 2002-010132/01.

XX Vaccine containing overlapping Th-cytotoxic T lymphocyte epitopes,  
 PT useful for protecting against human immune deficiency virus -  
 XX Example 1; Page 6; 11pp; English.

XX The invention relates to a vaccine comprising a mixture, or linear array,  
 CC of specific peptides that are human immunodeficiency virus (HIV-1) Th  
 CC (immunodominant T-helper epitope)-cytotoxic T lymphocyte (CTL), or their  
 CC variants. Peptides are also described which contain SIV (simian  
 CC immunodeficiency virus) derived epitopes. The vaccine is used to protect  
 CC against infection by HIV. The vaccines are based on an analysis of HLA  
 CC (human leukocyte antigen) alleles present in subject populations targeted  
 CC for vaccination and the commonest variants of HIV in the locality. They  
 CC contain sufficient immunogenic epitopes to ensure effective presentation  
 CC by almost all members of the target population. The present sequence is  
 CC an HIV-1 epitope based prototype vaccine of the invention.

XX Sequence 24 AA;

Query Match 100.0%; Score 48; DB 23; Length 24;

Best Local Similarity 100.0%; Pred. No. 0.64; Mismatches 0; Gaps 0;

Matches 7; Conservative 0; Indels 0; Gaps 0;

```

RESULT 11
AAU12479
ID AAU12479 standard; Peptide: 32 AA.
XX
AC AAU12479;
XX
DT 27-SEP-2001 (first entry)
XX
DE HIV Th-CTL peptide prototype vaccine immunogen #16.
XX
KW Human immunodeficiency virus; HIV; immunogen; vaccine; anti-HIV;
KW human leukocyte antigen; HLA; T-helper epitope; CTL; Th; epitope;
KW major histocompatibility complex; MHC; cytotoxic T-lymphocyte;
KW vaccinia ankara.
XX
OS Homo sapiens.
XX
PN MO200156355-A2.
XX
PD 09-AUG-2001.
XX
PF 05-FEB-2001; 2001WO-US03540.
XX
PR 04-FEB-2000; 2000US-0497497.
XX
PA (UYDU-) UNIV DUKE.
XX
PI Haynes BF, Liao H;
XX
DR WPI; 2001-488827/53.
XX
PT Vaccine for immunizing against human immunodeficiency virus has mixture
PT or linear array of peptides comprising immunodominant T-helper epitopes
PT and major histocompatibility complex cytotoxic T-lymphocyte epitopes -
XX
PS Claim 1; Page 24; 33pp; English.
XX
CC The present invention relates to human immunodeficiency virus (HIV)
CC and in particular to a human leukocyte antigen (HLA)-based HIV vaccine.
CC The vaccine comprises a mixture or linear array of peptides, or its
CC variants, where the peptides contain immunodominant T-helper (Th)
CC epitopes and major histocompatibility complex (MHC) cytotoxic T-
CC lymphocyte (CTL) epitopes and the linear array of peptides are
CC preferably expressed in modified vaccinia ankara. The vaccine is useful
CC for immunizing a patient against HIV and focuses immune response on many
CC dominant and subdominant CTL epitopes of HIV. DNA or live vectors with
CC linear arrays of CTL epitopes can be used as either primes or boosts of
CC peptides or of each other to optimally give CTL anti-HIV responses. The
CC vaccine induces salutary anti-HIV immune responses. AAU12447-AAU12558
CC represent the amino acid sequences of the Th-CTL epitopes and HIV
CC immunogenic peptides used in the invention.
XX
SQ Sequence 32 AA;
XX
Query Match 100.0%; Score 48; DB 22; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LMNWFNI 7
| | | | |
Db 3 LMNWFNI 9
XX
RESULT 12
AAU70242
ID AAU70242 standard; Peptide: 32 AA.
XX
AC AAU70242;
XX
DT 14-FEB-2002 (first entry)
XX
DE HIV Th-CTL overlapping epitopes prototype vaccine D-Th/D-CTL.
XX

```

```

KW HIV; human immunodeficiency virus; AIDS; STV; anti-HIV; virucide;
KW acquired immunodeficiency syndrome; simian immunodeficiency virus;
KW vaccine; cytotoxic T cell response inducer; Th; human leukocyte antigen;
KW immunodominant T-helper epitope; cytotoxic T lymphocyte; CTL; HLA.
XX
OS Human immunodeficiency virus type 1.
XX
OS Synthetic.
XX
PN US2001036461-A1.
XX
PD 01-NOV-2001.
XX
PF 05-FEB-2001; 2001US-0775805.
XX
PR 04-FEB-2000; 2000US-0497497.
XX
PA (HAYN/) HAYNES B F.
XX
PA (LIAO/) LIAO H.
XX
PI Haynes BF, Liao H;
XX
DR WPI; 2002-010132/01.
XX
PT Vaccine containing overlapping Th-cytotoxic T lymphocyte epitopes,
PT useful for protecting against human immune deficiency virus -
XX
PS Claim 1; Page 6; 11pp; English.
XX
CC The invention relates to a vaccine comprising a mixture, or linear array,
CC of specific peptides that are human immunodeficiency virus (HIV-1) Th
CC (immunodominant T-helper epitope)-cytotoxic T lymphocyte (CTL), or their
CC variants. Peptides are also described which contain STV (simian
CC immunodeficiency virus) derived epitopes. The vaccine is used to protect
CC against infection by HIV. The vaccines are based on an analysis of HLA
CC (human leukocyte antigen) alleles present in subject populations targeted
CC for vaccination and the commonest variants of HIV in the locality. They
CC contain sufficient immunogenic epitopes to ensure effective presentation
CC by almost all members of the target population. The present sequence is
CC an HIV-1 epitope based prototype vaccine of the invention.
XX
SQ Sequence 32 AA;
XX
Query Match 100.0%; Score 48; DB 23; Length 32;
Best Local Similarity 100.0%; Pred. No. 0.86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 LMNWFNI 7
| | | | |
Db 3 LMNWFNI 9
XX
RESULT 13
ABB09199
ID ABB09199 standard; Peptide: 33 AA.
XX
AC ABB09199;
XX
DT 04-JUL-2002 (first entry)
XX
DE HIV-1 gp160 amino acid sequence.
XX
KW HIV-1; human immunodeficiency virus; gp41; virology; vaccine.
XX
OS Human immunodeficiency virus type 1.
XX
PN RU2179980-C2.
XX
PD 27-FEB-2002.
XX
PF 20-MAR-2000; 2000RU-0106709.
XX
PR 20-MAR-2000; 2000RU-0106709.
XX

```

PA (VECT-) VECTOR VIROLOGY & BIOTECHN RES CENTRE.  
 PA (BIOM-) BIOMEDITSINSKII TSENTR NON COMMERCE ORGA.  
 XX  
 PI Tumanova Yu O, Kuvshinov VN, Melamed NV, Ushakova TA;  
 PI Masharskii Eh A, Illichev AA, Klimov NA, Kozlov AP, Sandakhchev LS;  
 XX  
 DR WPI: 2002-302480/34.  
 XX  
 PT Peptide-simulator of human immunodeficiency type-1 virus protein gp41  
 PT conservative epitope recognized by virus-neutralizing monoclonal  
 PT antibody 2f5 (variants) -  
 XX  
 PS Disclosure; Fig 3; 6pp; Russian.  
 XX  
 CC The present invention describes 12 peptide-simulators (ABB09187 to  
 CC ABB09198) of HIV-1 (human immunodeficiency virus type 1) protein gp41  
 CC conservative epitope prepared by affinity selection from phage peptide  
 CC libraries. These peptides differ from the HIV-1 protein gp41 conservative  
 CC epitope sequence but retain the ability to bind with the virus-  
 CC neutralising monoclonal antibodies 2F5. The invention also describes a  
 CC vaccine against HIV based on the peptides. The peptides can be used in  
 CC biotechnology, virology, and vaccines. The present sequence represents  
 CC an HIV-1 gp160 amino acid sequence given in the exemplification of the  
 CC present invention.  
 CC  
 SQ Sequence 33 AA;  
 XX  
 Query Match 100.0%; Score 48; DB 23; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 0.89;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 LMNMENI 7  
 |||||||  
 Db 18 LMNMENI 24  
 XX  
 RESULT 14  
 AAM17022  
 ID AAM17022 standard; peptide; 36 AA.  
 XX  
 AC AAM17022;  
 XX  
 DT 01-JUL-1997 (first entry)  
 XX  
 DE DP-178-derived peptide, T636, having high antiviral activity.  
 XX  
 KM HIV; SIV; human; simian immunodeficiency virus; glycoprotein 41;  
 KM transmembrane; antiviral; DP-178; DP-107; vaccine; inhibition;  
 KW replication; transmission.  
 XX  
 OS Synthetic.  
 XX  
 PN WO9640191-A1.  
 XX  
 PD 19-DEC-1996.  
 XX  
 PF 06-JUN-1996; 96WO-US09499.  
 XX  
 PR 07-JUN-1995; 95US-0481957.  
 XX  
 PA (TRIM-) TRIMERIS INC.  
 XX  
 PI Johnson RM, Lambert DM;  
 XX  
 DR WPI: 1997-099886/09.  
 XX  
 PT Compens. contg. DP-178 or DP-107 in combination with other  
 PT therapeutic agent - useful for treatment of HIV infection, esp. by  
 PT inhibiting replication or transmission of HIV.  
 XX  
 PS Claim 8; Figure 5A (III-IV); 84pp; English.  
 XX  
 CC AAM17021-WI7028 are DP-178-derived peptides that have high antiviral

CC activity. DP-178 is a peptide corresponding to residues 638-673 of HIV  
 CC type 1 glycoprotein 41 (gp41) transmembrane protein. DP-178 and its  
 CC derivatives/homologues are used in combination with a therapeutic agent,  
 CC e.g. a reverse transcriptase, viral protease, cytokine, glycosylation or  
 CC viral mRNA processing inhibitor or a nucleoside inhibitor. The peptides  
 CC work by inhibiting viral replication or inhibiting transmission. They  
 CC may also be used in vaccines for protecting against HIV infection.  
 CC  
 SQ Sequence 36 AA;  
 XX  
 Query Match 100.0%; Score 48; DB 18; Length 36;  
 Best Local Similarity 100.0%; Pred. No. 0.97;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 LMNMENI 7  
 |||||||  
 Db 28 LMNMENI 34  
 XX  
 RESULT 15  
 AAY89188  
 ID AAY89188 standard; peptide; 36 AA.  
 XX  
 AC AAY89188;  
 XX  
 DT 23-MAY-2000 (first entry)  
 XX  
 DE Core polypeptide fragment T No. 627.  
 XX  
 KM Retrovirus; hybrid polypeptide; enhancer; gp41; envelope protein; HIV-1;  
 KM HIV-2; SIV; pharmacokinetic; half-life; growth factor; cytokine; viral;  
 KM anti-fusogenic; differentiation factor; interleukin; interferon;  
 KM colony stimulating factor; hormone; angiogenic factor.  
 XX  
 OS Unidentified.  
 XX  
 PN WO959615-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 20-MAY-1999; 99WO-US11219.  
 XX  
 PR 20-MAY-1998; 98US-0082279.  
 XX  
 PA (TRIM-) TRIMERIS INC.  
 XX  
 PI Barney S, Guthrie KI, Merutka G, Anwer MK, Lambert DM;  
 XX  
 DR WPI: 2000-136792/12.  
 XX  
 PT A new hybrid polypeptide with enhanced pharmacokinetic properties  
 PT comprises enhancer sequence -  
 XX  
 PS Disclosure; Page 31; 124pp; English.  
 XX  
 CC The invention relates to hybrid polypeptides comprising enhancer peptide  
 CC sequence linked to core polypeptides. The enhancer polypeptides are  
 CC derived from various retroviral envelope (gp41) protein sequences,  
 CC especially from HIV-1, HIV-2 and SIV. The enhancer peptides enhance the  
 CC pharmacokinetic properties such as increasing the half-life of any core  
 CC polypeptide that they are linked to. The core polypeptides are any  
 CC polypeptide that may be introduced into a living system and that can  
 CC function as a pharmacologically useful peptide for the treatment or  
 CC prevention of a disease. The core polypeptides are bioactive peptides  
 CC selected from a growth factor, cytokine, differentiation factor,  
 CC interleukin, interferon, colony stimulating factor, hormone or  
 CC angiogenic factor. The peptides of the invention can be used for  
 CC inhibiting viral infection and can be used in anti-viral and  
 CC anti-fusogenic treatments. Sequences AAY8651-Y90055 represent core  
 CC polypeptide fragments that can be used in the invention. Some sequences  
 CC among those indicated also comprise enhancer fragments at terminal ends  
 CC and form hybrid polypeptides.

SO Sequence 36 AA;

Query Match 100.0%; Score 48; DB 21; Length 36;

Best Local Similarity 100.0%; Pred. No. 0.97;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNWFNI 7

|||||

Db 19 LMNWFNI 25

Search completed: March 18, 2003, 14:33:26  
job time : 6 secs



GenCore version 5.1.4\_p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:33:36 ; Search time 1.68966 Seconds  
(without alignments)  
121.895 Million cell updates/sec

Title: US-09-877-606-7  
Perfect score: 48  
Sequence: 1 LMNMFNI 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PTCUTS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*  
  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	20	1 US-08-218-025A-7	Sequence 7, Appli
2	48	100.0	23	1 US-08-218-025A-152	Sequence 152, App
3	48	100.0	36	3 US-08-484-223B-232	Sequence 232, App
4	48	100.0	36	3 US-08-484-223B-233	Sequence 233, App
5	48	100.0	36	4 US-09-082-279B-549	Sequence 549, App
6	48	100.0	36	4 US-09-082-279B-550	Sequence 550, App
7	48	100.0	36	4 US-09-082-279B-551	Sequence 551, App
8	48	100.0	36	4 US-09-082-279B-552	Sequence 552, App
9	48	100.0	36	4 US-09-082-279B-553	Sequence 553, App
10	48	100.0	36	4 US-09-082-279B-554	Sequence 554, App
11	48	100.0	36	4 US-09-082-279B-555	Sequence 555, App
12	48	100.0	36	4 US-09-082-279B-556	Sequence 556, App
13	48	100.0	36	4 US-09-082-279B-557	Sequence 557, App
14	48	100.0	36	4 US-09-082-279B-558	Sequence 558, App
15	48	100.0	36	4 US-09-082-279B-559	Sequence 559, App
16	48	100.0	36	4 US-09-082-279B-560	Sequence 560, App
17	48	100.0	36	4 US-08-474-349A-401	Sequence 401, App
18	48	100.0	36	4 US-08-474-349A-402	Sequence 402, App
19	48	100.0	36	4 US-08-474-349A-403	Sequence 403, App
20	48	100.0	36	4 US-08-474-349A-404	Sequence 404, App
21	48	100.0	36	4 US-08-474-349A-405	Sequence 405, App
22	48	100.0	36	4 US-08-474-349A-406	Sequence 406, App
23	48	100.0	36	4 US-08-474-349A-407	Sequence 407, App
24	48	100.0	36	4 US-08-474-349A-408	Sequence 408, App
25	48	100.0	36	4 US-08-474-349A-409	Sequence 409, App
26	48	100.0	36	4 US-08-474-349A-410	Sequence 410, App
27	48	100.0	36	4 US-08-474-349A-411	Sequence 411, App

28	48	100.0	36	4 US-08-474-349A-412	Sequence 412, App
29	48	100.0	36	4 US-09-315-304B-549	Sequence 549, App
30	48	100.0	36	4 US-09-315-304B-550	Sequence 550, App
31	48	100.0	36	4 US-09-315-304B-551	Sequence 551, App
32	48	100.0	36	4 US-09-315-304B-552	Sequence 552, App
33	48	100.0	36	4 US-09-315-304B-553	Sequence 553, App
34	48	100.0	36	4 US-09-315-304B-554	Sequence 554, App
35	48	100.0	36	4 US-09-315-304B-555	Sequence 555, App
36	48	100.0	36	4 US-09-315-304B-556	Sequence 556, App
37	48	100.0	36	4 US-09-315-304B-557	Sequence 557, App
38	48	100.0	36	4 US-09-315-304B-558	Sequence 558, App
39	48	100.0	36	4 US-09-315-304B-559	Sequence 559, App
40	48	100.0	36	4 US-09-315-304B-560	Sequence 560, App
41	48	100.0	38	4 US-09-082-279B-776	Sequence 776, App
42	48	100.0	38	4 US-09-315-304B-776	Sequence 776, App
43	48	100.0	39	4 US-09-082-279B-777	Sequence 777, App
44	48	100.0	39	4 US-09-315-304B-777	Sequence 777, App
45	48	100.0	40	4 US-09-082-279B-778	Sequence 778, App

## ALIGNMENTS

RESULT 1  
: Sequence 7, Application US/08218025A  
: Patent No. 5556744  
: GENERAL INFORMATION:  
: APPLICANT: Weiner, David B.  
: APPLICANT: Ugen, Kenneth E.  
: APPLICANT: Williams, William V.  
: TITLE OF INVENTION: Methods and Compositions for Diagnosing  
: TITLE OF INVENTION: and Treating Certain HIV Infected Patients  
: NUMBER OF SEQUENCES: 197  
: CORRESPONDENCE ADDRESSES:  
: ADDRESSSEE: Howson and Howson  
: STREET: P.O. Box 457, 321 No. 55567441stowen Road  
: CITY: Spring House  
: STATE: Pennsylvania  
: COUNTRY: U.S.A.  
: ZIP: 19477  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.25  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/08/218,025A  
: FILING DATE: 24-MAR-1994  
: CLASSIFICATION: 424  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 07/891,451  
: FILING DATE: 29-MAY-1992  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Bak, Mary E.  
: REGISTRATION NUMBER: 31,215  
: REFERENCE/DOCKET NUMBER: WSP33A  
: TELEPHONE: (215) 540-9206  
: TELEFAX: (215) 540-5818  
: INFORMATION FOR SEQ ID NO: 7:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 20 amino acids  
: TYPE: amino acid  
: TOPOLOGY: unknown  
: MOLECULE TYPE: peptide  
: US-08-218-025A-7  
  
Query Match 100.0% Score 48; DB 1; Length 20;  
Best Local Similarity 100.0% Pred. No. 0.11;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMNMFNI 7

Db 8 LMWMEI 14

RESULT 2  
US-08-218-025A-152

; Sequence 152, Application US/08218025A  
; Patent No. 5556744  
; GENERAL INFORMATION:  
; APPLICANT: Weiner, David B.  
; APPLICANT: Ugen, Kenneth E.  
; APPLICANT: Williams, William V.  
; TITLE OF INVENTION: Methods and Compositions for Diagnosing  
; TITLE OF INVENTION: and Treating Certain HIV Infected Patients  
; NUMBER OF SEQUENCES: 197  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Howson and Howson  
; STREET: P.O. Box 457, 321 No. 55567441stow Road  
; CITY: Spring House  
; STATE: Pennsylvania  
; COUNTRY: U.S.A.  
; ZIP: 19477

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/218,025A  
; FILING DATE: 24-MAR-1994  
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/891,451  
; FILING DATE: 29-MAY-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bak, Mary E.  
; REGISTRATION NUMBER: 31,215  
; REFERENCE/DOCKET NUMBER: NST33A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (215) 540-9206  
; TELEFAX: (215) 540-5818

; INFORMATION FOR SEQ ID NO: 152:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 23 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-218-025A-152

Query Match 100.0%; Score 48; DB 1; Length 23;  
Best Local Similarity 100.0%; Pred. No. 0.13; 0; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

QY 1 LMWMEI 7  
Db 9 LMWMEI 15

RESULT 3  
US-08-484-223B-232

; Sequence 232, Application US/08484223B  
; Patent No. 6020459  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Danl P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Peteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; TITLE OF INVENTION: TRANSMISSION

; NUMBER OF SEQUENCES: 245  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,223B  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:  
; NAME: Coituzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 7872-029  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 232:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 36 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown

; MOLECULE TYPE: peptide  
US-08-484-223B-232

Query Match 100.0%; Score 48; DB 3; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMWMEI 7  
Db 30 LMWMEI 36

RESULT 4  
US-08-484-223B-233

; Sequence 233, Application US/08484223B  
; Patent No. 6020459  
; GENERAL INFORMATION:  
; APPLICANT: Bolognesi, Danl P.  
; APPLICANT: Matthews, Thomas J.  
; APPLICANT: Wild, Carl T.  
; APPLICANT: Barney, Shawn O.  
; APPLICANT: Lambert, Dennis M.  
; APPLICANT: Peteway, Stephen R.  
; APPLICANT: Langlois, Alphonse J.  
; TITLE OF INVENTION: COMPOSITIONS FOR INHIBITION OF  
; TITLE OF INVENTION: MEMBRANE FUSION-ASSOCIATED EVENTS, INCLUDING HIV  
; TITLE OF INVENTION: TRANSMISSION

; NUMBER OF SEQUENCES: 245  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036-2711

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,223B

; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-9741/8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 233:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 36 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-484-223B-233

Query Match 100.0%; Score 48; DB 3; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
|||||  
Db 29 LMNMFNI 35

RESULT 5  
US-09-082-279B-549  
; Sequence 549, Application US/09082279B  
; Patent No. 6258782  
; GENERAL INFORMATION:  
; APPLICANT: Barney, Shawn  
; APPLICANT: Guthrie, Kelly  
; APPLICANT: Merutka, Gene  
; APPLICANT: Anwer, Mohamed  
; APPLICANT: Lambert, Dennis  
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
; FILE REFERENCE: 7872-043  
; CURRENT APPLICATION NUMBER: US/09/082,279B  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 1515  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 549  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Core polypeptide  
US-09-082-279B-549

Query Match 100.0%; Score 48; DB 4; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
|||||  
Db 19 LMNMFNI 25

RESULT 6  
US-09-082-279B-550  
; Sequence 550, Application US/09082279B  
; Patent No. 6258782  
; GENERAL INFORMATION:  
; APPLICANT: Barney, Shawn  
; APPLICANT: Guthrie, Kelly  
; APPLICANT: Merutka, Gene  
; APPLICANT: Anwer, Mohamed  
; APPLICANT: Lambert, Dennis  
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED

; TITLE OF INVENTION: PHARMACOKINETIC PROPERTIES  
; FILE REFERENCE: 7872-043  
; CURRENT APPLICATION NUMBER: US/09/082,279B  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 1515  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 550  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Core polypeptide  
US-09-082-279B-550

Query Match 100.0%; Score 48; DB 4; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
|||||  
Db 20 LMNMFNI 26

RESULT 7  
US-09-082-279B-551  
; Sequence 551, Application US/09082279B  
; Patent No. 6258782  
; GENERAL INFORMATION:  
; APPLICANT: Barney, Shawn  
; APPLICANT: Guthrie, Kelly  
; APPLICANT: Merutka, Gene  
; APPLICANT: Anwer, Mohamed  
; APPLICANT: Lambert, Dennis  
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
; FILE REFERENCE: 7872-043  
; CURRENT APPLICATION NUMBER: US/09/082,279B  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 1515  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 551  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Core polypeptide  
US-09-082-279B-551

Query Match 100.0%; Score 48; DB 4; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
|||||  
Db 21 LMNMFNI 27

RESULT 8  
US-09-082-279B-552  
; Sequence 552, Application US/09082279B  
; Patent No. 6258782  
; GENERAL INFORMATION:  
; APPLICANT: Barney, Shawn  
; APPLICANT: Guthrie, Kelly  
; APPLICANT: Merutka, Gene  
; APPLICANT: Anwer, Mohamed  
; APPLICANT: Lambert, Dennis  
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
; FILE REFERENCE: 7872-043  
; CURRENT APPLICATION NUMBER: US/09/082,279B  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 1515

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 552
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-552

Query Match          100.0%; Score 48; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7
Db 22 LMNMFNI 28

RESULT 9
US-09-082-279B-553
; Sequence 553, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 553
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-553

Query Match          100.0%; Score 48; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7
Db 23 LMNMFNI 29

RESULT 10
US-09-082-279B-554
; Sequence 554, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 554
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

```

; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-554

Query Match          100.0%; Score 48; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7
Db 24 LMNMFNI 30

RESULT 11
US-09-082-279B-555
; Sequence 555, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 555
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-555

Query Match          100.0%; Score 48; DB 4; Length 36;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7
Db 25 LMNMFNI 31

RESULT 12
US-09-082-279B-556
; Sequence 556, Application US/09082279B
; Patent No. 6258782
; GENERAL INFORMATION:
; APPLICANT: Barney, Shawn
; APPLICANT: Guthrie, Kelly
; APPLICANT: Merutka, Gene
; APPLICANT: Anwer, Mohamed
; APPLICANT: Lambert, Dennis
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED
; FILE REFERENCE: 7872-043
; CURRENT APPLICATION NUMBER: US/09/082,279B
; CURRENT FILING DATE: 1998-05-20
; NUMBER OF SEQ ID NOS: 1515
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 556
; LENGTH: 36
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Core polypeptide
US-09-082-279B-556

Query Match          100.0%; Score 48; DB 4; Length 36;
```

Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
Db 26 LMNMFNI 32

RESULT 13  
US-09-082-279B-557  
; Sequence 557, Application US/09082279B  
; Patent No. 6258782  
; GENERAL INFORMATION:  
; APPLICANT: Barney, Shawn  
; APPLICANT: Guthrie, Kelly  
; APPLICANT: Merutka, Gene  
; APPLICANT: Anwer, Mohamed  
; APPLICANT: Lambert, Dennis  
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
; FILE REFERENCE: 7872-043  
; CURRENT APPLICATION NUMBER: US/09/082,279B  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 1515  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 557  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Core polypeptide  
US-09-082-279B-557

Query Match 100.0%; Score 48; DB 4; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
Db 27 LMNMFNI 33

RESULT 14  
US-09-082-279B-558  
; Sequence 558, Application US/09082279B  
; Patent No. 6258782  
; GENERAL INFORMATION:  
; APPLICANT: Barney, Shawn  
; APPLICANT: Guthrie, Kelly  
; APPLICANT: Merutka, Gene  
; APPLICANT: Anwer, Mohamed  
; APPLICANT: Lambert, Dennis  
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
; FILE REFERENCE: 7872-043  
; CURRENT APPLICATION NUMBER: US/09/082,279B  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 1515  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 558  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Core polypeptide  
US-09-082-279B-558

Query Match 100.0%; Score 48; DB 4; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
Db 27 LMNMFNI 33

Db 28 LMNMFNI 34

RESULT 15  
US-09-082-279B-559  
; Sequence 559, Application US/09082279B  
; Patent No. 6258782  
; GENERAL INFORMATION:  
; APPLICANT: Barney, Shawn  
; APPLICANT: Guthrie, Kelly  
; APPLICANT: Merutka, Gene  
; APPLICANT: Anwer, Mohamed  
; APPLICANT: Lambert, Dennis  
; TITLE OF INVENTION: HYBRID POLYPEPTIDES WITH ENHANCED  
; FILE REFERENCE: 7872-043  
; CURRENT APPLICATION NUMBER: US/09/082,279B  
; CURRENT FILING DATE: 1998-05-20  
; NUMBER OF SEQ ID NOS: 1515  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 559  
; LENGTH: 36  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Core polypeptide  
US-09-082-279B-559

Query Match 100.0%; Score 48; DB 4; Length 36;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
Db 29 LMNMFNI 35

Search completed: March 18, 2003, 14:50:05  
Job time : 1.68966 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.4.p5\_4578  
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:36:16 ; Search time 1.62069 Seconds  
(without alignments)  
199.079 Million cell updates/sec

Title: US-09-877-606-7

Perfect score: 48

Sequence: 1 LMNMFNI 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues

Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/PCT05\_PUBCOMB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	56	10 US-09-779-451-4	Sequence 4, Appl1
2	48	100.0	268	10 US-09-854-816-9	Sequence 9, Appl1
3	48	100.0	268	10 US-09-854-816-10	Sequence 10, Appl1
4	48	100.0	268	10 US-09-854-816-16	Sequence 16, Appl1
5	48	100.0	268	10 US-09-854-816-17	Sequence 17, Appl1
6	48	100.0	268	10 US-09-854-816-18	Sequence 18, Appl1
7	48	100.0	268	10 US-09-854-816-19	Sequence 19, Appl1
8	48	100.0	268	10 US-09-854-816-41	Sequence 41, Appl1
9	48	100.0	269	10 US-09-854-816-12	Sequence 12, Appl1
10	48	100.0	269	10 US-09-854-816-31	Sequence 31, Appl1
11	48	100.0	269	10 US-09-854-816-43	Sequence 43, Appl1
12	48	100.0	269	10 US-09-854-816-44	Sequence 44, Appl1
13	48	100.0	269	10 US-09-854-816-46	Sequence 46, Appl1
14	48	100.0	345	10 US-09-779-451-8	Sequence 8, Appl1
15	48	100.0	519	10 US-09-756-551A-8	Sequence 8, Appl1
16	48	100.0	844	10 US-09-991-258-19	Sequence 19, Appl1
17	48	100.0	853	9 US-10-003-035-33	Sequence 33, Appl1
18	48	100.0	856	10 US-09-476-242-1	Sequence 1, Appl1
19	48	100.0	1101	9 US-10-003-035-53	Sequence 53, Appl1

20	48	100.0	1186	9 US-10-003-035-55	Sequence 55, Appl1
21	44	91.7	37	10 US-09-779-451-68	Sequence 68, Appl1
22	44	91.7	47	10 US-09-779-451-66	Sequence 66, Appl1
23	43	89.6	145	12 US-10-000-321-11	Sequence 11, Appl1
24	43	89.6	146	12 US-10-000-321-10	Sequence 10, Appl1
25	43	89.6	233	10 US-09-854-816-50	Sequence 50, Appl1
26	43	89.6	233	10 US-09-854-816-51	Sequence 51, Appl1
27	43	89.6	267	10 US-09-854-816-15	Sequence 15, Appl1
28	43	89.6	268	10 US-09-854-816-8	Sequence 8, Appl1
29	43	89.6	268	10 US-09-854-816-13	Sequence 13, Appl1
30	43	89.6	268	10 US-09-854-816-14	Sequence 14, Appl1
31	43	89.6	268	10 US-09-854-816-26	Sequence 26, Appl1
32	43	89.6	268	10 US-09-854-816-35	Sequence 35, Appl1
33	43	89.6	268	10 US-09-854-816-39	Sequence 39, Appl1
34	43	89.6	268	10 US-09-854-816-40	Sequence 40, Appl1
35	43	89.6	268	10 US-09-854-816-68	Sequence 68, Appl1
36	43	89.6	268	10 US-09-854-816-70	Sequence 70, Appl1
37	43	89.6	268	10 US-09-854-816-71	Sequence 71, Appl1
38	43	89.6	268	10 US-09-854-816-72	Sequence 72, Appl1
39	43	89.6	269	10 US-09-854-816-20	Sequence 20, Appl1
40	43	89.6	269	10 US-09-854-816-21	Sequence 21, Appl1
41	43	89.6	269	10 US-09-854-816-22	Sequence 22, Appl1
42	43	89.6	269	10 US-09-854-816-24	Sequence 24, Appl1
43	43	89.6	269	10 US-09-854-816-28	Sequence 28, Appl1
44	43	89.6	269	10 US-09-854-816-30	Sequence 30, Appl1
45	43	89.6	269	10 US-09-854-816-33	Sequence 33, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-779-451-4  
Sequence 4, Application US/09779451  
Patent No. US20020094521A1  
GENERAL INFORMATION:  
APPLICANT: Wild, Carl T.  
TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors  
FILE REFERENCE: 1900.0300003  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: US 60/235,901  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/181,543  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Human Immunodeficiency virus type 1  
US-09-779-451-4

Query Match  
Best Local Similarity 100.0%; Score 48; DB 10; Length 56;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7  
|||||||  
DB 47 LMNMFNI 53

RESULT 2  
US-09-854-816-9  
Sequence 9, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Pheasant  
Melissa A. Starovaanik

```
James A. Wells
TITLE OF INVENTION: Constrained Helical Peptides and Methods of
Making Same
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/965,056
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Ph.D., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1005R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 9:
US-09-854-816-9

Query Match      100.0%; Score 48; DB 10; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7
    |||||
Db 199 LMNMFNI 205

RESULT 3
US-09-854-816-10
Sequence 10, Application US/09854816
Patent No. US20020151473A1
GENERAL INFORMATION:
APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovasmnik
James A. Wells
TITLE OF INVENTION: Constrained Helical Peptides and Methods of
Making Same
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
```

```
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/965,056
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Ph.D., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1005R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 268 amino acids
TYPE: Amino Acid
TOPOLOGY: Linear
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-09-854-816-10

Query Match      100.0%; Score 48; DB 10; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNMFNI 7
    |||||
Db 199 LMNMFNI 205

RESULT 4
US-09-854-816-16
Sequence 16, Application US/09854816
Patent No. US20020151473A1
GENERAL INFORMATION:
APPLICANT: Andrew C. Braisted
J. Kevin Judice
Robert S. McDowell
J. Christopher Phelan
Melissa A. Starovasmnik
James A. Wells
TITLE OF INVENTION: Constrained Helical Peptides and Methods of
Making Same
NUMBER OF SEQUENCES: 113
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genentech, Inc.
STREET: 1 DNA Way
CITY: South San Francisco
STATE: California
COUNTRY: USA
ZIP: 94080
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Winpatin (Genentech)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/854,816
FILING DATE: 15-May-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/965,056
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Torchia, Ph.D., Timothy E.
REGISTRATION NUMBER: 36,700
REFERENCE/DOCKET NUMBER: P1005R2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650/225-8674
TELEFAX: 650/952-9881
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
```



LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-854-816-16

Query Match 100.0%; Score 48; DB 10; Length 268;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNMWFI 7  
|||||  
DB 199 LNMWFI 205

RESULT 5  
US-09-854-816-17  
Sequence 17, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-854-816-17

Query Match 100.0%; Score 48; DB 10; Length 268;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNMWFI 7  
|||||  
DB 199 LNMWFI 205

RESULT 6

US-09-854-816-18  
Sequence 18, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-854-816-18

Query Match 100.0%; Score 48; DB 10; Length 268;  
Best Local Similarity 100.0%; Pred. No. 3.1;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNMWFI 7  
|||||  
DB 199 LNMWFI 205

RESULT 7  
US-09-854-816-19  
Sequence 19, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way

```

? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Winpatin (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/854,816
? FILING DATE: 15-May-2001
? CLASSIFICATION: <Unknown>
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 08/965,056
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Torchia, Phd., Timothy E.
? REGISTRATION NUMBER: 36,700
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-8674
? TELEFAX: 650/952-9881
? INFORMATION FOR SEQ ID NO: 19:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 268 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
? SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-854-816-19

Query Match          100.0%; Score 48; DB 10; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7
Db 199 LMNWFNI 205

RESULT 8
US-09-854-816-41
? Sequence 41, Application US/09854816
? Patent No. US20020151473A1
? GENERAL INFORMATION:
? APPLICANT: Andrew C. Braisted
? J. Kevin Judice
? Robert S. McDowell
? J. Christopher Phelan
? Melissa A. Starovasnuk
? James A. Wells
?
? TITLE OF INVENTION: Constrained Helical Peptides and Methods of
? Making Same
?
? NUMBER OF SEQUENCES: 113
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 1 DNA Way
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Winpatin (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/854,816
? FILING DATE: 15-May-2001
? CLASSIFICATION: <Unknown>
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 08/965,056
? FILING DATE: <Unknown>
?
? 3
```

```

? ATTORNEY/AGENT INFORMATION:
? NAME: Torchia, Phd., Timothy E.
? REGISTRATION NUMBER: 36,700
? REFERENCE//DOCKET NUMBER: P1005R2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-8674
? TELEFAX: 650/952-9881
? INFORMATION FOR SEQ ID NO: 41:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 268 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
? SEQUENCE DESCRIPTION: SEQ ID NO: 41:
US-09-854-816-41

Query Match          100.0%; Score 48; DB 10; Length 268;
Best Local Similarity 100.0%; Pred. No. 3.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7
Db 199 LMNWFNI 205

RESULT 9
US-09-854-816-12
? Sequence 12, Application US/09854816
? Patent No. US20020151473A1
? GENERAL INFORMATION:
? APPLICANT: Andrew C. Braisted
? J. Kevin Judice
? Robert S. McDowell
? J. Christopher Phelan
? Melissa A. Starovasnuk
? James A. Wells
?
? TITLE OF INVENTION: Constrained Helical Peptides and Methods of
? Making Same
?
? NUMBER OF SEQUENCES: 113
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Genentech, Inc.
? STREET: 1 DNA Way
? CITY: South San Francisco
? STATE: California
? COUNTRY: USA
? ZIP: 94080
?
? COMPUTER READABLE FORM:
? MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Winpatin (Genentech)
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/854,816
? FILING DATE: 15-May-2001
? CLASSIFICATION: <Unknown>
? PRIORITY APPLICATION DATA:
? APPLICATION NUMBER: 08/965,056
? FILING DATE: <Unknown>
? ATTORNEY/AGENT INFORMATION:
? NAME: Torchia, Phd., Timothy E.
? REGISTRATION NUMBER: 36,700
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 650/225-8674
? TELEFAX: 650/952-9881
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 269 amino acids
? TYPE: Amino Acid
? TOPOLOGY: Linear
? SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-854-816-12

Query Match          100.0%; Score 48; DB 10; Length 269;
```

Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNMWFI 7  
Db 200 LNMWFI 206

## RESULT 10

US-09-854-816-31  
; Sequence 31, Application US/09854816  
; Patent No. US20020151473A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrew C. Braisted  
; J. Kevin Judice  
; Robert S. McDowell  
; J. Christopher Phelan  
; Melissa A. Starovasnik  
; James A. Wells  
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
; Making Same  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/854,816  
; FILING DATE: 15-May-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/965,056  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Torchia, Ph.D., Timothy E.  
; REGISTRATION NUMBER: 36,700  
; REFERENCE/DOCKET NUMBER: P1005R2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-8674  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 31:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 269 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 31:  
US-09-854-816-31  
Query Match 100.0%; Score 48; DB 10; Length 269;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LNMWFI 7  
Db 200 LNMWFI 206  
RESULT 11  
US-09-854-816-43  
; Sequence 43, Application US/09854816  
; Patent No. US20020151473A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrew C. Braisted  
; J. Kevin Judice  
; Robert S. McDowell

J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
Making Same

NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/854,816

FILING DATE: 15-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Ph.D., Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P1005R2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 269 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 43:

US-09-854-816-43  
Query Match 100.0%; Score 48; DB 10; Length 269;  
Best Local Similarity 100.0%; Pred. No. 3.2;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNMWFI 7  
Db 200 LNMWFI 206

## RESULT 12

US-09-854-816-44  
; Sequence 44, Application US/09854816  
; Patent No. US20020151473A1  
; GENERAL INFORMATION:  
; APPLICANT: Andrew C. Braisted  
; J. Kevin Judice  
; Robert S. McDowell  
; J. Christopher Phelan  
; Melissa A. Starovasnik  
; James A. Wells  
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
; Making Same  
; NUMBER OF SEQUENCES: 113  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 44:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 44:
US-09-854-816-44

Query Match      100.0%; Score 48; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LMNWFNI 7
        |||||
Db      200 LMNWFNI 206

RESULT 13
US-09-854-816-46
; Sequence 46, Application US/09854816
; Patent No. US20020151473A1
; GENERAL INFORMATION:
; APPLICANT: Andrew C. Braisted
; J. Kevin Judice
; Robert S. McDowell
; J. Christopher Phelan
; Melissa A. Starovasnik
; James A. Wells
; TITLE OF INVENTION: Constrained Helical Peptides and Methods of
; Making Same
; NUMBER OF SEQUENCES: 113
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/854,816
; FILING DATE: 15-May-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/965,056
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Torchia, PhD., Timothy E.
; REGISTRATION NUMBER: 36,700
; REFERENCE/DOCKET NUMBER: P1005R2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-8674
; TELEFAX: 650/952-9881
```

```

; INFORMATION FOR SEQ ID NO: 46:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 269 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 46:
US-09-854-816-46

Query Match      100.0%; Score 48; DB 10; Length 269;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LMNWFNI 7
        |||||
Db      200 LMNWFNI 206

RESULT 14
US-09-779-451-8
; Sequence 8, Application US/09779451
; Patent No. US20020094521A1
; GENERAL INFORMATION:
; APPLICANT: Wild, Carl T.
; TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors
; FILE REFERENCE: 1900.0300003
; CURRENT APPLICATION NUMBER: US/09/779,451
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 60/235,901
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/181,543
; PRIOR FILING DATE: 2000-02-10
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 8
; LENGTH: 345
; TYPE: PROT
; ORGANISM: Human Immunodeficiency virus type 1
US-09-779-451-8

Query Match      100.0%; Score 48; DB 10; Length 345;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 LMNWFNI 7
        |||||
Db      158 LMNWFNI 164

RESULT 15
US-09-756-551A-8
; Sequence 8, Application US/09756551A
; Patent No. US20020051768A1
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
; NUCLEIC ACID AND METHODS OF MAKING AND
; USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,551A
```

CC -!- FUNCTION: Inhibits spontaneous contractions in several muscle  
CC groups. May be involved in the expansion phase of feeding  
CC behaviour in sea anemones.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- TISSUE SPECIFICITY: Neuron-specific.  
DR InterPro: IPR001023; Hsp70.  
KW Neuropeptide; Amidation.  
FT CHAIN 1 4 ANTHO-RIAMIDE I.  
FT CHAIN 2 4 ANTHO-RIAMIDE II.  
FT MOD\_RES 1 1 L-3-PHENYLACTYL.  
FT MOD\_RES 4 4 AMIDATION.  
SQ SEQUENCE 4 AA; 598 MW; 60441B59A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
Db 1 F 1

## RESULT 15

OCPL\_OCTMI STANDARD; PRT; 4 AA.  
AC P58648;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Cardioactive peptides Ocp-1/Ocp-2.  
OS Octopus minor (Octopus).  
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;  
OC Incurrata; Octopodidae; Octopus.  
OX NCBI\_TaxID=89766;  
RN [1]  
RP SEQUENCE, SYNTHESIS, MASS SPECTROMETRY, AND CHARACTERIZATION.  
RC TISSUE-Brain;  
RA MEDLINE-20336815; PubMed-10876044;  
RA Iwakoshi E., Hisada M., Minakata H.;  
RT "Cardioactive peptides isolated from the brain of a Japanese octopus,  
RT Octopus minor."  
RL Peptides 21:623-630(2000).  
CC -!- FUNCTION: Cardioactive; has both positive chronotropic and  
CC inotropic effects on the heart. Ocp-2 is a 1000 time less  
CC active than Ocp-1.  
CC -!- SUBCELLULAR LOCATION: Secreted.  
CC -!- PTM: Ocp-2 has L-Phe instead of D-Phe.  
CC -!- MASS SPECTROMETRY: MW=395.2; METHOD-MALDI.  
KW Hormone; D-amino acid.  
FT MOD\_RES 2 2 D-PHENYLANILINE.  
SQ SEQUENCE 4 AA; 394 MW; 6AA879C810000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
Db 1 G 1

Search completed: March 18, 2003, 14:34:06  
Job time : 4 secs

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
Db 1 F 1

## RESULT 12

FLRN\_ANTLR  
ID FLRN\_ANTLR STANDARD; PRT; 4 AA.  
AC P58707;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE Antho-RNamide.  
DE Anthopleura elegantissima (Sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;  
OC Myantheae; Actiniidae; Anthopleura.  
OX NCBI\_TaxID=6110;  
RN [1]  
RP SEQUENCE, AND MASS-SPECTROMETRY.  
RX MEDLINE=90319122; PubMed=1973541;  
RA Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,  
RA Reinscheid R.K., Notackner H.-P., Staley A.L.;  
RT Isolation of L-3-phenylacetyl-Leu-Arg-Asn-NH<sub>2</sub> (Antho-RNamide), a sea  
RT anemone neuropeptide containing an unusual amino-terminal blocking  
RT group.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- TISSUE SPECIFICITY: Neuron-specific.  
CC -1- MASS SPECTROMETRY: MW=549.3; METHOD=FAB.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 1 1 L-3-PHENYLACTYL.  
FT MOD\_RES 4 4 AMIDATION.  
SQ SEQUENCE 4 AA; 549 MW; 64540729A0000000 CRC64;  
Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 X 1  
Db 1 F 1

## RESULT 13

FMRF\_MACNI STANDARD; PRT; 4 AA.  
ID FMRF\_MACNI  
AC P01162;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DE 15-JUN-1998 (Rel. 36, Last annotation update)  
DE FMRFamide (Peak C) (Cardioexcitatory neuropeptide).  
OS Macrocaltista nimbosa (Sun-ray clam),  
OS Nerereis virens (Sandworm),  
OS Hirudo medicinalis (Medical leech), and  
OS Helisoma trivolvis (Snail).  
OC Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroida;  
OC Veneroida; Veneridae; Macrocallista.  
OX NCBI\_TaxID=6594, 6353, 6421, 27815;  
RN [1]  
RP SEQUENCE, AND SYNTHESIS.  
RC SPECIES=M.nimbosa; TISSUE=Cerebral pedal, and Visceral ganglion;  
RX MEDLINE=77215956; PubMed=877582;  
RA Price D.A., Greenberg M.J.;  
RT Structure of a molluscan cardioexcitatory neuropeptide.";  
RL Science 197:670-671(1977).  
RN [2]  
RP SEQUENCE, AND CHARACTERIZATION.  
RC SPECIES=M.nimbosa; TISSUE=Ganglion;  
RX MEDLINE=78012038; PubMed=909875;  
RA Price D.A., Greenberg M.J.;

RT "Purification and characterization of a cardioexcitatory neuropeptide  
RT from the central ganglia of a bivalve mollusc.";  
RL Prep. Biochem. 7:261-281(1977).  
RN [3]  
RP SEQUENCE.  
RC SPECIES=N.virens;  
RX MEDLINE=90259866; PubMed=2342992;  
RA Krajinak K.G., Price D.A.;  
RT "Authentic FMRFamide is present in the polychaete Nerereis virens.";  
RL Peptides 11:75-77(1990).  
RN [4]  
RP SEQUENCE.  
RC SPECIES=H.medicalis;  
RX MEDLINE=92195954; PubMed=1686933;  
RA Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;  
RT "Identification of Rnamide neuropeptides in the medicinal leech.";  
RL Peptides 12:897-908(1991).  
RN [5]  
RP SEQUENCE.  
RC SPECIES=H.trivolvis; TISSUE=Kidney;  
RX MEDLINE=94286417; PubMed=7912428;  
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;  
RT "FMRFamide-related peptides from the kidney of the snail, Helisoma  
RT trivolvis.";  
RL Peptides 15:31-36(1994).  
CC -1- FUNCTION: MYOACTIVE; CARDIOEXCITATORY SUBSTANCE. PHARMACOLOGICAL  
CC ACTIVITIES INCLUDE AUGMENTATION, INDUCTION, AND REGULATION OF  
CC CARDIAC CONTRACTION.  
CC -1- SIMILARITY: BELONGS TO THE FMRF (FMRFAMIDE RELATED PEPTIDE)  
CC FAMILY.  
CC PIR: A01426; ECKN.  
DR PIR: A60418; A60418.  
KW Neuropeptide; Amidation.  
FT MOD\_RES 4 4 AMIDATION.  
SQ SEQUENCE 4 AA; 600 MW; 69D40699A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
Db 1 F 1

## RESULT 14

FYRI\_ANTLR STANDARD; PRT; 4 AA.  
ID FYRI\_ANTLR  
AC P58706;  
DT 15-JUN-2002 (Rel. 41, Created)  
DT 15-JUN-2002 (Rel. 41, Last sequence update)  
DE 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Antho-RNamide I [contains: Antho-RNamide II].  
OS Anthopleura elegantissima (Sea anemone).  
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;  
OC Myantheae; Actiniidae; Anthopleura.  
OX NCBI\_TaxID=6110;  
RN [1]  
RP SEQUENCE.  
RX MEDLINE=92270459; PubMed=1821096;  
RA Notackner H.-P., Rinehart K.L. Jr., McFarlane I.D.,  
RA Grimmelikhuijzen C.J.P.;  
RT "Isolation of two novel neuropeptides from sea anemones: the unusual,  
RT biologically active L-3-phenylacetyl-Tyr-Arg-Ile-NH<sub>2</sub> and its  
RT des-phenylacetyl fragment Tyr-Arg-Ile-NH<sub>2</sub>.";  
RL Peptides 12:1165-1173(1991).  
RN [2]  
RP FUNCTION.  
RX MEDLINE=93391436; PubMed=8397415;  
RA McFarlane I.D., Hudman D., Notackner H.-P., Grimmelikhuijzen C.J.P.;  
RT "The expansion behaviour of sea anemones may be coordinated by two  
RT inhibitory neuropeptides, Antho-RNamide and Antho-RNamide.";  
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).

```

RESULT 8
FAR3_HIRME STANDARD: PRT: 4 AA.
AC P42562;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FMRamide-like neuropeptide YLRF-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
OC Arynchobdellida; Hirudiniiformes; Hirudinae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsotis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -I- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD.RES 4 4
SQ SEQUENCE 4 AA; 598 MW; 69D4073B30000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 Y 1

RESULT 9
FAR4_HIRME STANDARD: PRT: 4 AA.
AC P42563;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FMRamide-like neuropeptide YMRF-amide.
OS Hirudo medicinalis (Medicinal leech).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
OC Arynchobdellida; Hirudiniiformes; Hirudinae; Hirudo.
OX NCBI_TaxID=6421;
RN [1]
RP SEQUENCE.
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsotis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
CC -I- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD.RES 4 4
SQ SEQUENCE 4 AA; 616 MW; 69D4068B30000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 1 Y 1

RESULT 10
FFKA_ATEL STANDARD: PRT: 4 AA.
AC P58705;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Antho-Kamide.
OS Anthopleura elegantissima (Sea anemone).
OC Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actinaria;
OC Nyantheae; Actiniidae; Anthopleura.
OX NCBI_TaxID=6110;
RN [1]
RP SEQUENCE.
RX MEDLINE=92028652; PubMed=1681803;
RA Notackner H.-P., Rinehart K.L. Jr., Grimmelikhuijzen C.J.P.;
RT "Isolation of L-3-phenylactyl-Phe-Lys-Ala-NH2 (Antho-Kamide), a
RT novel neuropeptide from sea anemones.";
RL Biochem. Biophys. Res. Commun. 179:1205-1211(1991).
RN [2]
RP FUNCTION.
RX MEDLINE=93391436; PubMed=8397415;
RA McFarlane I.D., Hudson D., Notackner H.-P., Grimmelikhuijzen C.J.P.;
RT "The expansion behaviour of sea anemones may be coordinated by two
RT inhibitory neuropeptides, Antho-Kamide and Antho-Ramide.";
RL Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC -I- FUNCTION: Inhibits spontaneous contractions in several muscle
CC groups. May be involved in the expansion phase of feeding
CC behaviour in sea anemones.
CC -I- SUBCELLULAR LOCATION: Secreted.
CC -I- TISSUE SPECIFICITY: Neuron-specific.
KW Neuropeptide; Amidation.
FT MOD.RES 1 1
FT MOD.RES 4 4
SQ SEQUENCE 4 AA; 512 MW; 6DD339C9A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1
DB 4 A 4

RESULT 11
FLRF_HIRME STANDARD: PRT: 4 AA.
AC P42561;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE FLRamide.
OS Hirudo medicinalis (Medicinal leech), and
OS Helisoma trivolvis (Snail).
OC Eukaryota; Metazoa; Annelida; Clitellata; Hirudinea; Hirudinea;
OC Arynchobdellida; Hirudiniiformes; Hirudinae; Hirudo.
OX NCBI_TaxID=6421, 27815;
RN [1]
RP SEQUENCE.
RX SPECIES-H. medicinalis;
RX MEDLINE=92195954; PubMed=1686933;
RA Evans B.D., Pohl J., Kartsotis M.A., Calabrese R.L.;
RT "Identification of Rfamide neuropeptides in the medicinal leech.";
RL Peptides 12:897-908(1991).
RN [2]
RP SEQUENCE.
RX SPECIES-H. trivolvis;
RX MEDLINE=94286417; PubMed=7912428;
RA Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT "FMRamide-related peptides from the kidney of the snail, Helisoma
RT trivolvis.";
RL Peptides 15:31-36(1994).
CC -I- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC FAMILY.
KW Neuropeptide; Amidation.
FT MOD.RES 4 4
SQ SEQUENCE 4 AA; 582 MW; 69D40729A0000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 4;

```

RA Iwashita T., Nomoto K.;  
 RT "Crystal structure and molecular conformation of achatin-1  
 (H-Gly-D-Phe-Ala-Asp-OH), an endogenous neuropeptide containing a  
 RT D-amino acid residue.";  
 RL Int. J. Pept. Protein Res. 39:258-264(1992).  
 CC -I- FUNCTION: NEUROEXCITATORY PEPTIDE; INCREASES THE IMPULSE FREQUENCY  
 CC AND PRODUCES A SPIKE BROADENING OF THE IDENTIFIED HEART EXCITATORY  
 CC NEURON (PON); ALSO ENHANCES THE AMPLITUDE AND FREQUENCY OF THE  
 CC HEART BEAT. HAS ALSO AN EFFECT ON SEVERAL OTHER MUSCLES.  
 DR PIR: A32480; A32480.  
 KW Hormone; D-amino acid.  
 FT MOD\_RES 2  
 SQ SEQUENCE 4 AA; 408 MW; 6AADD9C810000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 DB 3 A 3  
 RESULT 5  
 DCML\_PSECH STANDARD; PRT; 4 AA.  
 AC P19916;  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Carbon monoxide dehydrogenase large chain (EC 1.2.99.2) (CO  
 DE dehydrogenase subunit L) (CO-DH L) (Fragment).  
 GN CUNL.  
 OS Pseudomonas carboxydohydrogena.  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Bradyrhizobium group.  
 OX NCBI\_TaxID=290;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=90055678; PubMed=2818128;  
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;  
 RT "Homology and distribution of CO dehydrogenase structural genes in  
 RT carboxydohydrophic bacteria.";  
 RL Arch. Microbiol. 152:335-341(1989).  
 CC -I- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon  
 CC dioxide.  
 CC -I- CATALYTIC ACTIVITY: CO + H(2)O + acceptor -> CO(2) + reduced  
 CC acceptor.  
 CC -I- COFACTOR: Molybdenum (molybdopterin).  
 CC -I- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND  
 CC SMALL.  
 DR PIR: P10140; P10140.  
 DR Oxidoreductase; Molybdenum.  
 KW NON\_TER 4  
 SQ SEQUENCE 4 AA; 441 MW; 7761EB76F0000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 DB 1 M 1  
 RESULT 6  
 DCMS\_PSECH STANDARD; PRT; 4 AA.  
 AC P19918;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 15-JUN-2002 (Rel. 41, Last annotation update)  
 DE Carbon monoxide dehydrogenase small chain (EC 1.2.99.2) (CO

DE dehydrogenase subunit S) (CO-DH S) (Fragment).  
 GN CUNL.  
 OS Pseudomonas carboxydohydrogena.  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Bradyrhizobium group.  
 OX NCBI\_TaxID=290;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=90055678; PubMed=2818128;  
 RA Kraut M., Hugendieck I., Herwig S., Meyer O.;  
 RT "Homology and distribution of CO dehydrogenase structural genes in  
 RT carboxydohydrophic bacteria.";  
 RL Arch. Microbiol. 152:335-341(1989).  
 CC -I- FUNCTION: Catalyzes the oxidation of carbon monoxide to carbon  
 CC dioxide.  
 CC -I- CATALYTIC ACTIVITY: CO + H(2)O + acceptor -> CO(2) + reduced  
 CC acceptor.  
 CC -I- COFACTOR: BINDS TWO 2FE-2S CLUSTERS.  
 CC -I- SUBUNIT: CONSISTS OF THREE POLYPEPTIDE CHAINS: LARGE, MEDIUM, AND  
 CC SMALL.  
 DR PIR: P10146; P10146.  
 DR Oxidoreductase; Iron-sulfur.  
 KW NON\_TER 4  
 SQ SEQUENCE 4 AA; 420 MW; 6DD3DD6F00000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 DB 2 A 2  
 RESULT 7  
 EOSL\_HUMAN STANDARD; PRT; 4 AA.  
 ID EOSL\_HUMAN  
 AC P02731;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last annotation update)  
 DE Eosinophilolactac peptide.  
 DE Eosinophilolactac peptides.  
 OS Homo sapiens (Human).  
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE.  
 RX MEDLINE=76078412; PubMed=1060093;  
 RA Goetzl E.J., Austen K.F.;  
 RT "Purification and synthesis of eosinophilolactac tetrapeptides of  
 RT human lung tissue: identification as eosinophil chemotactic factor of  
 RT anaphylaxis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 72:4123-4127(1975).  
 CC -I- MISCELLANEOUS: THESE PEPTIDES ARE RELEASED FROM MAST CELLS IN LUNG  
 CC (AND OTHER TISSUES) DURING HYPERSENSITIVITY REACTIONS  
 CC (ANAPHYLAXIS). THEIR ACTIVITIES, PREFERENTIALLY AFFECTING  
 CC EOSINOPHILS, INCLUDE CHEMOTAXIS, CHEMOTACTIC DEACTIVATION, RELEASE  
 CC OF ENZYMES, AND STIMULATION OF THE HEXOSE MONOPHOSPHATE SHUNT.  
 DR PIR: A03190; ETHUL.  
 FT VARIANT 1  
 SQ SEQUENCE 4 AA; 390 MW; 6B05B862A0000000 CRC64;  
 Query Match 0.0%; Score 0; DB 1; Length 4;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 X 1  
 DB 3 S 3



CC -I- CATALYTIC ACTIVITY: ATP + an acid + protein -> AMP + diphosphate +  
CC an acyl-protein thioester.  
CC -I- PATHWAY: Bioluminescent fatty acid reduction system; second step.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: M62812; -; NOT\_ANNOTATED\_CDS.  
KW Bioluminescence; Ligase.  
FT NON\_TER 1  
SQ SEQUENCE 3 AA; 374 MW; 6AA3303000000000 CRC64;  
  
Query Match 0.0%; Score 0; DB 1; Length 3;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 x 1  
DB 1 1 1  
  
RESULT 3  
ID THYL\_PIG STANDARD; PRT; 3 AA.  
AC P01511;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Thyrolobberin (Thyrotropin releasing hormone) (TRH) (Protrellin).  
OS Sus scrofa (Pig).  
OS Ovis aries (Sheep).  
OS Bombina orientalis (Oriental fire-bellied toad), and  
OS Notophthalmus viridescens (Eastern newt) (Triturus viridescens).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suiua; Suidae; Sus.  
OX NCBI\_TaxID=9823, 9940, 8346, 8316;  
RN [1]  
RP SEQUENCE.  
RC SPECIES=Pig; TISSUE=Hypothalamus;  
RX MEDLINE=70136150; PubMed=4984938;  
RA Nair R.M.G., Barrett J.F., Bowers C.Y., Schally A.V.;  
RT "Structure of porcine thyrotropin releasing hormone.";  
RL Biochemistry 9:1103-1106(1970).  
RN [2]  
RP SYNTHESIS.  
RC SPECIES=Pig;  
RX MEDLINE=70039904; PubMed=4982117;  
RA Bolter J., Enzmann F., Folkers K., Bowers C.Y., Schally A.V.;  
RT "The identity of chemical and hormonal properties of the thyrotropin  
RT releasing hormone and pyroglutamy-histidyl-proline amide.";  
RL Biochem. Biophys. Res. Commun. 37:705-710(1969).  
RN [3]  
RP SEQUENCE.  
RC SPECIES=Sheep; TISSUE=Hypothalamus;  
RA Desiderio D.M. Jr., Burgess R., Dunn T.F., Vale W., Guillemin R.,  
RA Ward D.N.;  
RT "The elucidation of the primary structure of the hypothalamic thyroid  
RT stimulating hormone releasing factor of ovine origin by means of mass  
RT spectrometry.";  
RL Org. Mass Spectrom. 5:221-228(1971).  
RN [4]  
RP SYNTHESIS.  
RC SPECIES=Sheep;  
RX MEDLINE=70163386; PubMed=4985794;  
RA Burgess R., Dunn T.F., Desiderio D.M., Ward D.N., Vale W.,  
RA Guillemin R.;  
RT "Characterization of ovine hypothalamic hypophysiotropic  
RT TSH-releasing factor.";

RL Nature 226:321-325(1970).  
RN [5]  
RP SEQUENCE.  
RC SPECIES=B.orientalis; TISSUE=Skin;  
RX MEDLINE=76138399; PubMed=815011;  
RA Yasuhara T., Nakajima T.;  
RT "Letter: Occurrence of Pyl-His-Pro-NH2 in the frog skin.";  
RL Letten. Pharm. Bull. 23:3301-3303(1975).  
RN [6]  
RP SEQUENCE.  
RC SPECIES=N.viridescens;  
RX MEDLINE=75035605; PubMed=4214528;  
RA Grilme-Joergensen Y., McKelvy J.F.;  
RT "Biosynthesis of thyrotropin releasing factor by newt (Triturus  
RT viridescens) brain in vitro. Isolation and characterization of  
RT thyrotropin releasing factor.";  
RL J. Neurochem. 23:471-478(1974).  
CC -I- FUNCTION: TRH FUNCTIONS AS A REGULATOR OF THE BIOSYNTHESIS OF TSH  
CC IN THE ANTERIOR PITUITARY GLAND AND AS A NEUROTRANSMITTER/  
CC NEUROMODULATOR IN THE CENTRAL AND PERIPHERAL NERVOUS SYSTEMS.  
CC  
CC PIR: A01415; RHPGT.  
DR PIR: A93750; RSHST.  
DR PIR: A90919; RHTDPO.  
DR PIR: A92971; A92971.  
KM Amidation.  
FT MOD\_RES 1 1 PYROLIDONE CARBOXYLIC ACID.  
FT MOD\_RES 3 3 AMIDATION.  
SQ SEQUENCE 3 AA; 380 MW; 7761F6B000000000 CRC64;  
  
Query Match 0.0%; Score 0; DB 1; Length 3;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 x 1  
DB 1 Q 1  
  
RESULT 4  
ID ACH1\_ACHFU STANDARD; PRT; 4 AA.  
AC P35904;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Achatin-I.  
OS Achatina fulica (Giant African snail).  
OC Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;  
OC Achatinacea; Achatinidae; Achatina.  
OX NCBI\_TaxID=6530;  
RN [1]  
RP SEQUENCE, CHARACTERIZATION, AND SYNTHESIS.  
RC STRAIN=Ferussac; TISSUE=Ganglion;  
RX MEDLINE=89273551; PubMed=2597281;  
RA Kamatani Y., Minakata H., Kenny P.T.M., Iwashita T., Watanabe K.,  
RA Funase K., Sun X.P., Yongsiri A., Kim K.H., Novales-Li P.,  
RA Novales E.T., Kanapi C.G., Takeuchi H., Nomoto K.;  
RT "Achatin-I, an endogenous neuroexcitatory tetrapeptide from Achatina  
RT fulica Ferussac containing a D-amino acid residue.";  
RL Biochem. Biophys. Res. Commun. 160:1015-1020(1989).  
RN [2]  
RP CHARACTERIZATION.  
RC STRAIN=Ferussac; TISSUE=Heart atrium;  
RX MEDLINE=91264856; PubMed=1675568;  
RA Fujimoto K., Kubota I., Yasuda-Kamatani Y., Minakata H., Nomoto K.,  
RA Yoshida M., Harada A., Muneoka Y., Kobayashi M.;  
RT "Purification of achatin-I from the atria of the African giant snail,  
RT Achatina fulica, and its possible function.";  
RL Biochem. Biophys. Res. Commun. 177:847-853(1991).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY.  
RX MEDLINE=93014529; PubMed=1399265;  
RA Ishida T., In Y., Doi M., Inoue M., Yasuda-Kamatani Y., Minakata H.,

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:27:26 ; Search time 1 Seconds  
(without alignments)

290.334 Million cell updates/sec

Title: US-09-877-606-8

Sequence: 1 xxxxxxx 7

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	3	1 GRWM_HUMAN	P01157 homo sapien
2	0	0.0	3	1 LUXE_VIBFI	P24272 vibrio fisc
3	0	0.0	3	1 THYL_PIG	P01151 sus scrofa
4	0	0.0	4	1 ACHI_ACHFU	P35904 achatina fu
5	0	0.0	4	1 DCML_PSECH	P19916 pseudomonas
6	0	0.0	4	1 DCMS_PSECH	P19918 pseudomonas
7	0	0.0	4	1 EOST_HUMAN	P02731 homo sapien
8	0	0.0	4	1 FAR3_HIRME	P42562 hirudo medi
9	0	0.0	4	1 FAR4_HIRME	P42563 hirudo medi
10	0	0.0	4	1 FEKA_AMEL	P58705 anthopleura
11	0	0.0	4	1 FLRF_HIRME	P42561 hirudo medi
12	0	0.0	4	1 FLRN_AMEL	P58707 anthopleura
13	0	0.0	4	1 FMRE_MACNI	P01162 macrocallis
14	0	0.0	4	1 FYRI_AMEL	P58706 anthopleura
15	0	0.0	4	1 OCPI_OCTMI	P58648 octopus min
16	0	0.0	4	1 OCP3_OCTMI	P58649 octopus min
17	0	0.0	4	1 RM01_YEAST	P36515 saccharomyc
18	0	0.0	4	1 TUFT_HUMAN	P01856 homo sapien
19	0	0.0	5	1 ALI4_CARMA	P81817 carclinus ma
20	0	0.0	5	1 BIOA_CIFER	P13071 citrobacter
21	0	0.0	5	1 B10B_CIFER	P12997 citrobacter
22	0	0.0	5	1 BPP7_BOTIN	P30425 bothrops in
23	0	0.0	5	1 E103_LITRU	P82099 litorea rub
24	0	0.0	5	1 E104_LITRU	P82100 litorea rub
25	0	0.0	5	1 FARP_ARTTR	P41853 attiposthi
26	0	0.0	5	1 PAP2_PARMA	P81864 pardachilus
27	0	0.0	5	1 PRC7_PERAM	P01373 periplaneta
28	0	0.0	5	1 PSK_DAUCA	P58261 dancus caro
29	0	0.0	5	1 RE11_LITRU	P82070 litorea rub
30	0	0.0	5	1 RE21_LITRU	P82071 litorea rub
31	0	0.0	5	1 RE31_LITRU	P82072 litorea rub
32	0	0.0	5	1 RE32_LITRU	P82073 litorea rub
33	0	0.0	5	1 SUGA_ACHDO	P19991 acheta dome

34	0	0.0	5	1 TPIS_CANEA	P54714 canis famli
35	0	0.0	5	1 TRM3_ECOLI	P13973 escherichia
36	0	0.0	5	1 UC22_MAIZE	P80628 zea mays (m
37	0	0.0	5	1 UP01_MOUSE	P38639 mus musculu
38	0	0.0	5	1 UXA4_CHLTR	P38003 chlamydia t
39	0	0.0	6	1 ACPH_RABYT	P25154 oryctolagus
40	0	0.0	6	1 ASP2_LACSN	P82655 lactobacilli
41	0	0.0	6	1 CIP1_MYTED	P13736 mytilus edu
42	0	0.0	6	1 CIP2_MYTED	P13737 mytilus edu
43	0	0.0	6	1 E101_LITRU	P82096 litorea rub
44	0	0.0	6	1 FARP_MONEX	P41966 moniezia ex
45	0	0.0	6	1 LOK1_LOCMI	P41491 locusta mlg

## ALIGNMENTS

RESULT 1  
GRWM\_HUMAN STANDARD; PRT; 3 AA.  
ID GRWM\_HUMAN  
AC P01157;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 21-JUL-1986 (Rel. 01, Last annotation update)  
DE Growth-modulating peptide.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN (1)  
RP SEQUENCE.  
RX MEDLINE=77162369; PubMed=858356;  
RA Schlesinger D.H., Pickart L., Thaler M.M.;  
RT "Growth-modulating serum tripeptide is glycyL-histidyl-L-lysine.";  
RL Experientia 33:324-325(1977).  
CC -I- MISCELLANEOUS: THIS SERUM TRIPEPTIDE HAS BEEN FOUND TO STIMULATE  
CC GROWTH OF SOME CELL TYPES AND TO INHIBIT OTHER TYPES IN VITRO.  
DR PIR: A01421; GRNU.  
SO SEQUENCE 3 AA; 340 MW; 6331EB1000000000 CRC64;

Query Match 0.0%; Score 0; DB 1; Length 3;  
Best local similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
DB 1 G 1  
RESULT 2  
LUXE\_VIBFI STANDARD; PRT; 3 AA.  
ID LUXE\_VIBFI  
AC P24272;  
DT 01-MAR-1992 (Rel. 21, Created)  
DT 01-MAR-1992 (Rel. 21, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Long-chain-fatty-acid--luciferin-component ligase (EC 6.2.1.19) (acyl-  
DE protein synthetase) (Fragment).  
GN LUXE.  
OS Vibrio fischeri.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=668;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91072226; PubMed=2254256;  
RA Swartzman E., Kapoor S., Graham A.F., Meighen E.A.;  
RT "A new Vibrio fischeri lux gene precedes a bidirectional termination  
RT site for the lux operon.";  
RL J. Bacteriol. 172:6797-6802(1990).  
CC -I- FUNCTION: ACYL-PROTEIN SYNTHETASE ACTIVATES TETRADECANOIC ACID.  
CC IT IS A COMPONENT OF THE FATTY ACID REDUCTASE COMPLEX RESPONSIBLE  
CC FOR CONVERTING TETRADECANOIC ACID TO THE ALDEHYDE WHICH SERVES AS  
CC SUBSTRATE IN THE LUCIFERASE-CATALYZED REACTION.

QY 1 X 1  
Db 1 Q 1

## RESULT 14

E37196  
bradykinin-potentiating peptide 5 - island jararaca  
C:Species: Bothrops insularis (island jararaca)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: E37196  
R:Cintrà, A.C.O.; Vieira, C.A.; Giglio, J.R.  
J. Protein Chem. 9, 221-227, 1990  
A:Title: Primary structure and biological activity of bradykinin potentiating peptides  
A:Reference number: A37196; MUID:90351557; PMID:2386615  
A:Accession: E37196  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <CIN>  
C:Keywords: pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

## RESULT 15

E37196  
bradykinin-potentiating peptide 6 - island jararaca  
C:Species: Bothrops insularis (island jararaca)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: E37196  
R:Cintrà, A.C.O.; Vieira, C.A.; Giglio, J.R.  
J. Protein Chem. 9, 221-227, 1990  
A:Title: Primary structure and biological activity of bradykinin potentiating peptides  
A:Reference number: A37196; MUID:90351557; PMID:2386615  
A:Accession: E37196  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <CIN>  
C:Keywords: pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 Q 1

Search completed: March 18, 2003, 14:37:20  
Job time : 2.24138 secs

B23751  
spinal cord peptide SCP-5 - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: B23751  
R:Hi, K.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou, K.  
Arch. Biochem. Biophys. 240, 178-183, 1985  
A:Reference number: A23751; MUID:85250425; PMID:4015098  
A:Accession: B23751  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <HST>  
C:Superfamily: unassigned animal peptides

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 M 1

## RESULT 9

A33802  
thyrotropin-releasing hormone-like peptide - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A33802  
R:Cockle, S.M.; Aitken, A.; Beg, F.; Smyth, D.G.  
J. Biol. Chem. 264, 7786-7791, 1989  
A:Title: A novel peptide, pyroglutamylglutaminylproline amide, in the rabbit prostate comp

A:Reference number: A33802; MUID:89255196; PMID:2498305  
A:Accession: A33802  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-3 <COC>  
C:Superfamily: unassigned animal peptides  
C:Keywords: amidated carboxyl end; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 Q 1

## RESULT 10

A22565  
R-phycoerythrin alpha-1 chain - red alga (Gastrocionium coulteri) (fragment)  
C:Species: Gastrocionium coulteri  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A22565  
R:Klotz, A.V.; Glazer, A.N.  
J. Biol. Chem. 260, 4856-4863, 1985  
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.  
A:Reference number: A22565; MUID:85182601; PMID:3886644  
A:Accession: A22565  
A:Molecule type: protein  
A:Residues: 1-3 <KLO>

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 2 Y 2

RESULT 11  
PQ0010  
angiotensin-converting enzyme inhibitor (FLP-3) - common fig  
N:Alternate names: ficus latex peptide 3  
C:Species: Ficus carica (common fig)  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: PQ0010  
R:Maruyama, S.; Miyoshi, S.; Tanaka, H.  
Agric. Biol. Chem. 53, 2763-2767, 1989  
A:Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.  
A:Reference number: PQ0008  
A:Accession: PQ0010  
A:Molecule type: protein  
A:Residues: 1-3 <MAR>  
A:Experimental source: latex  
C:Keywords: angiotensin-converting enzyme inhibitor

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 L 1

## RESULT 12

S13894  
histidinol dehydrogenase (EC 1.1.1.23) - wild cabbage (fragment)  
C:Species: Brassica oleracea (wild cabbage)  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: S13894  
R:Nagal, A.; Scheidegger, A.  
Arch. Biochem. Biophys. 284, 127-132, 1991  
A:Title: Purification and characterization of histidinol dehydrogenase from cabbage.  
A:Reference number: S13894; MUID:91112783; PMID:1989490  
A:Accession: S13894  
A:Molecule type: protein  
A:Residues: 1-3 <NAG>  
A:Experimental source: var. capitata  
C:Keywords: dimer; NAD; oxidoreductase

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 X 1

## RESULT 13

A43391  
TRH-like tripeptide - alfalfa  
C:Species: Medicago sativa (alfalfa)  
C>Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A43391  
R:Jackey, D.B.  
J. Biol. Chem. 267, 17508-17511, 1992  
A:Title: Isolation and structural determination of a novel TRH-like tripeptide, pyrog  
A:Reference number: A43391; MUID:92388092; PMID:1517203  
A:Accession: A43391  
A:Molecule type: protein  
A:Residues: 1-3 <LAC>  
C:Keywords: amidated carboxyl end; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
DB 1 Q 1

## RESULT 3

RHSHT

thyroliberin - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)

C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A93750; A01415

R:Desiderio Jr., D.M.; Burgess, R.; Dunn, T.F.; Vale, W.; Guillemin, R.; Ward, D.N.

Org. Mass Spectrom. 5, 221-228, 1971

A:Title: The elucidation of the primary structure of the hypothalamic thyroïd stimulin

A:Reference number: A93750

A:Accession: A93750

A:Molecule type: protein

A:Residues: 1-3 &lt;DES&gt;

R:Burgess, R.; Dunn, T.F.; Desiderio, D.; Ward, D.N.; Vale, W.; Guillemin, R.

Nature 226, 321-325, 1970

A:Title: Characterization of ovine hypothalamic hypophysiotropic TSH-releasing factor.

A:Reference number: A93161; MUID:70163386; PMID:4985794

A:Contents: annotation

A:Note: Physicochemical characteristics and biological activities of the natural and syn

C:Keywords: thyroliberin precursor

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match  
Best Local Similarity 0.0%; Score 0; DB 3; Length 3;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 Q 1

## RESULT 4

thyroliberin - eastern newt (tentative sequence)

C:Species: Notophthalmus viridescens, Triturus viridescens (eastern newt)

C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A92971; A01415

R:Grimm-Jorgensen, Y.; McKelvy, J.F.

J. Neurochem. 23, 471-478, 1974

A:Title: Biosynthesis of thyrotropin releasing factor by newt (Triturus viridescens) bra

A:Reference number: A92971; MUID:75035605; PMID:4214528

A:Accession: A92971

A:Molecule type: protein

A:Residues: 1-3 &lt;GRIT&gt;

A:Note: a peptide with the chromatographic and electrophoretic characteristics of thyrol

ibidine, or glutamic acid

C:Superfamily: thyroliberin precursor

C:Keywords: amidated carboxyl end; cutaneous gland; hormone; hypothalamus; pyroglutamic

F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental

F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match  
Best Local Similarity 0.0%; Score 0; DB 3; Length 3;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 Q 1

## RESULT 5

GKHU

growth-modulating peptide - human

C:Species: Homo sapiens (man)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A01421

R:Schlesinger, D.H.; Pickart, L.; Thaler, M.M.

Experientia 33, 324-325, 1977

A:Title: Growth-modulating serum tripeptide is glycyl-histidyl-lysine.

A:Reference number: A01421; MUID:77162369; PMID:858356

A:Accession: A01421

A:Molecule type: protein

A:Residues: 1-3 &lt;SCH&gt;

A:Note: This serum tripeptide is found to stimulate growth of some cell types and to

C:Superfamily: unassigned animal peptides

Query Match  
Best Local Similarity 0.0%; Score 0; DB 3; Length 3;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 G 1

## RESULT 6

A60898

bursin - chicken

C:Species: Gallus gallus (chicken)

C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A60898

R:Audhya, T.; Kroon, D.; Heavner, G.; Viamontes, G.; Goldstein, G.

Science 231, 997-999, 1986

A:Title: Tripeptide structure of bursin, a selective B-cell-differentiating hormone o

A:Reference number: A60898; MUID:86122916; PMID:3484838

A:Accession: A60898

A:Molecule type: protein

A:Residues: 1-3 &lt;AUD&gt;

C:Superfamily: unassigned animal peptides

C:Keywords: amidated carboxyl end; hormone

F:3/Modified site: amidated carboxyl end (Gly) #status experimental

Query Match  
Best Local Similarity 0.0%; Score 0; DB 3; Length 3;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 K 1

## RESULT 7

A23751

spinal cord peptide SCP-4 - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001

C:Accession: A23751

R:Hsi, R.L.; Chen, R.L.; Chen, Z.G.; Zhang, H.L.; Lu, Y.A.; Guo, S.Y.; Wu, S.X.; Tsou

Arch. Biochem. Biophys. 240, 178-183, 1985

A:Reference number: A23751; MUID:85250425; PMID:4015098

A:Accession: A23751

A:Status: Preliminary

A:Molecule type: protein

A:Residues: 1-3 &lt;HSI&gt;

C:Superfamily: unassigned animal peptides

Query Match  
Best Local Similarity 0.0%; Score 0; DB 3; Length 3;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1

DB 1 Q 1

## RESULT 8

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:36 ; Search time 2.24138 Seconds  
(Without alignments)  
300.235 Million cell updates/sec

Title: US-09-877-606-8  
Perfect score: 7  
Sequence: 1 XXXXXX 7

Scoring table: BIOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	3	RHTDRO	thyroliberin - Bom
2	0	0.0	3	RHPGT	thyroliberin - pig
3	0	0.0	3	RHSMT	thyroliberin - she
4	0	0.0	3	A92971	thyroliberin - eas
5	0	0.0	3	GKHU	growth-modulating
6	0	0.0	3	A60898	burstin - chicken
7	0	0.0	3	A23751	spinal cord peptid
8	0	0.0	3	B23751	thyrotropin-releas
9	0	0.0	3	A33802	R-phycoerythrin al
10	0	0.0	3	A22565	angiotensin-conver
11	0	0.0	3	PQ0010	histidinol dehydro
12	0	0.0	3	S13894	TRH-like tripeptid
13	0	0.0	3	A43391	bradykinin-potenti
14	0	0.0	3	E37196	bradykinin-potenti
15	0	0.0	3	F37196	bradykinin-potenti
16	0	0.0	3	IS0412	gene p20K protein
17	0	0.0	3	PT0636	T-cell receptor be
18	0	0.0	3	PT0578	T-cell receptor be
19	0	0.0	3	PT0571	T-cell receptor be
20	0	0.0	3	PT0622	T-cell receptor be
21	0	0.0	3	178890	tyrosine protein k
22	0	0.0	3	S68328	blood cell protein
23	0	0.0	3	TI1892	cytochrome-c oxida
24	0	0.0	4	ECXAA	antho-Ramide neur
25	0	0.0	4	S18401	thyroglobulin - do
26	0	0.0	4	A02147	phagocytosis-stimu
27	0	0.0	4	A32039	tyrosine-melanocyt
28	0	0.0	4	ECNK	cardioexcitatory n
29	0	0.0	4	PL0140	carbon-monoxide de

30	0	0.0	4	2	PL0146	carbon-monoxide de
31	0	0.0	4	2	A37832	phenol 2-monooxyge
32	0	0.0	4	2	A48360	gamma subunit of p
33	0	0.0	4	2	I40697	biotin A - Citroba
34	0	0.0	4	2	A61300	22K superhelical D
35	0	0.0	4	2	I57745	D-mannanase hydrol
36	0	0.0	4	2	A41890	protein D - Escher
37	0	0.0	4	2	S43014	hypothetical prote
38	0	0.0	4	2	D41654	hypothetical prote
39	0	0.0	4	2	B43848	cell surface adhes
40	0	0.0	4	2	I40505	hypothetical prote
41	0	0.0	4	2	I40870	phospholipase C (E
42	0	0.0	4	2	I40804	endoglucanase F -
43	0	0.0	4	2	T46627	hypothetical prote
44	0	0.0	4	2	S09478	globulin IV alpha
45	0	0.0	4	2	S53508	starvation-induced

## ALIGNMENTS

RESULT 1  
RHTDRO  
thyroliberin - Bombina orientalis  
C:Species: Bombina orientalis  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A90919; A01415  
R:Yasuhara, T.; Nakajima, T.  
Chem. Pharm. Bull. 23, 3301-3303, 1975  
A:Title: Occurrence of Pyr-His-Pro-NH-2 in the frog skln.  
A:Reference number: A90919; MUID:76138399; PMID:815011  
A:Accession: A90919  
A:Molecule type: protein  
A:Residues: 1-3 <YAS>  
C:Superfamily: thyroliberin precursor  
C:Keywords: amidated carboxyl end; cutaneous gland; hormone; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 X 1  
Db 1 Q 1

RESULT 2  
RHPGT  
thyroliberin - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 15-Jun-2001 #sequence\_revision 15-Jun-2001 #text\_change 15-Jun-2001  
C:Accession: A01415  
R:Naier, R.M.G.; Barrelet, J.F.; Bowers, C.Y.; Schally, A.V.  
Biochemistry 9, 1103-1106, 1970  
A:Title: Structure of porcine thyrotropin releasing hormone.  
A:Reference number: A90560; MUID:70136150; PMID:4984938  
A:Accession: A01415  
A:Molecule type: protein  
A:Residues: 1-3 <NAI>  
R:Boiler, J.; Enzmann, F.; Folkers, K.; Bowers, C.Y.; Schally, A.V.  
Biochem. Biophys. Res. Commun. 37, 705-710, 1969  
A:Title: The identity of chemical and hormonal properties of the thyrotropin releasin  
A:Reference number: A90167; MUID:70039904; PMID:4982117  
A:Contents: annotation  
A:Note: Biological activities and Rf values (in 17 chromatographic systems) of the sy  
C:Superfamily: thyroliberin precursor  
C:Keywords: amidated carboxyl end; hormone; hypothalamus; pyroglutamic acid  
F:1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental  
F:3/Modified site: amidated carboxyl end (Pro) #status experimental

Query Match 0.0%; Score 0; DB 3; Length 3;

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 15:49:42 ; Search time 9.89655 Seconds

(without alignments)  
548.507 Million cell updates/sec

Title: US-09-877-606-8

Perfect score: 7

Sequence: 1 xxxxxxx 7

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 538826 segs, 387737923 residues

Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=x1h  
-O=/cg2\_1/USPTO.spool/US09877606/runat\_18032003\_134802\_17783/app\_query.fasta.1.1059  
-DB=Published.Applications.NA -OFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCU=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62  
-TRANS=human40.cdt -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0  
-MALEN=2000000000 -USER=US09877606.ecgn.1.1.115 @runat\_18032003\_134802\_17783  
-NCU=6 -ICPU=3 -NO\_XIPXY -NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT=LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA.\*

1: /cg2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cg2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cg2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cg2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cg2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cg2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cg2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cg2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cg2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cg2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cg2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cg2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cg2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cg2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	2	9	US-09-971-894-19
2	0	0.0	2	9	US-09-971-894-19
3	0	0.0	2	9	US-09-971-894-20
4	0	0.0	2	9	US-09-971-894-20

## ALIGNMENTS

5	0	0.0	2	10	US-09-735-363A-50	Sequence 50, Appl
6	0	0.0	2	10	US-09-735-363A-50	Sequence 50, Appl
7	0	0.0	2	10	US-09-735-363A-51	Sequence 51, Appl
8	0	0.0	2	10	US-09-735-363A-51	Sequence 51, Appl
9	0	0.0	3	9	US-09-796-692-3119	Sequence 3119, Ap
10	0	0.0	3	9	US-09-796-692-3119	Sequence 3119, Ap
11	0	0.0	3	9	US-09-971-894-21	Sequence 21, Appl
12	0	0.0	3	9	US-09-971-894-21	Sequence 21, Appl
13	0	0.0	3	9	US-09-971-894-22	Sequence 22, Appl
14	0	0.0	3	9	US-09-971-894-22	Sequence 22, Appl
15	0	0.0	3	9	US-10-127-645-1	Sequence 1, Appl
16	0	0.0	3	9	US-10-127-645-1	Sequence 1, Appl
17	0	0.0	3	10	US-09-735-363A-7	Sequence 7, Appl
18	0	0.0	3	10	US-09-735-363A-7	Sequence 7, Appl
19	0	0.0	3	10	US-09-735-363A-8	Sequence 8, Appl
20	0	0.0	3	10	US-09-735-363A-8	Sequence 8, Appl
21	0	0.0	3	10	US-09-804-653-11	Sequence 11, Appl
22	0	0.0	3	10	US-09-804-653-11	Sequence 11, Appl
23	0	0.0	3	10	US-09-816-763-10	Sequence 10, Appl
24	0	0.0	3	10	US-09-816-763-10	Sequence 10, Appl
25	0	0.0	4	9	US-09-942-087A-41	Sequence 41, Appl
26	0	0.0	4	9	US-09-942-087A-41	Sequence 41, Appl
27	0	0.0	4	9	US-09-942-087A-42	Sequence 42, Appl
28	0	0.0	4	9	US-09-942-087A-42	Sequence 42, Appl
29	0	0.0	4	9	US-09-942-087A-43	Sequence 43, Appl
30	0	0.0	4	9	US-09-942-087A-43	Sequence 43, Appl
31	0	0.0	4	9	US-09-992-028-3	Sequence 3, Appl
32	0	0.0	4	9	US-09-992-028-3	Sequence 3, Appl
33	0	0.0	4	9	US-10-098-514-18	Sequence 18, Appl
34	0	0.0	4	9	US-10-098-514-18	Sequence 18, Appl
35	0	0.0	4	9	US-10-117-641-4	Sequence 4, Appl
36	0	0.0	4	9	US-10-117-641-4	Sequence 4, Appl
37	0	0.0	4	9	US-10-117-641-8	Sequence 8, Appl
38	0	0.0	4	9	US-10-117-641-8	Sequence 8, Appl
39	0	0.0	4	9	US-10-117-641-9	Sequence 9, Appl
40	0	0.0	4	9	US-10-117-641-9	Sequence 9, Appl
41	0	0.0	4	9	US-09-835-370-53	Sequence 53, Appl
42	0	0.0	4	9	US-09-835-370-53	Sequence 53, Appl
43	0	0.0	4	9	US-10-179-053-1	Sequence 1, Appl
44	0	0.0	4	9	US-10-179-053-1	Sequence 1, Appl
45	0	0.0	4	9	US-09-911-020A-2	Sequence 2, Appl



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:42:33 ; Search time 5.37931 Seconds  
(without alignments)  
399.073 Million cell updates/sec

Title: US-09-877-606-8  
Perfect score: 7  
Sequence: 1 xxxxxxxx 7

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/uspro.spool/US09877606/runat\_18032003\_134801\_17715/app\_query.fasta\_1.1059  
-DB=Issued\_Patents\_NA -OPMT=fastact -SUFFIX=rni -MINMATCH=0.1 -LOOPEL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=plo -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09877606 @CGN.1.1.101 @runat\_18032003\_134801\_17715 -NCPu=6 -ICPU=3  
-NO\_XIPXY -NO\_MMAPP -LARGENOURRY -NEG\_SCORES=0 -NAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/1/ina/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCRNUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	4	US-09-177-650-79
2	0	0.0	1	4	US-09-177-650-79
3	0	0.0	1	5	PCT-US93-00977-702
4	0	0.0	1	5	PCT-US93-00977-702
5	0	0.0	2	1	US-08-268-679B-8
6	0	0.0	2	1	US-08-268-679B-8
7	0	0.0	2	1	US-08-268-679B-8
8	0	0.0	2	1	US-08-457-274A-16
9	0	0.0	2	1	US-08-457-274A-16
10	0	0.0	2	1	US-08-484-192-16
11	0	0.0	2	1	US-08-484-192-16
12	0	0.0	2	2	US-08-726-464B-28

13	0	0.0	2	3	US-09-016-520-35	Sequence 35, Appl
14	0	0.0	2	3	US-09-016-520-35	Sequence 35, Appl
15	0	0.0	2	4	US-09-130-973-35	Sequence 35, Appl
16	0	0.0	2	4	US-09-130-973-35	Sequence 35, Appl
17	0	0.0	2	4	US-09-477-902-35	Sequence 35, Appl
18	0	0.0	2	4	US-09-477-902-35	Sequence 35, Appl
19	0	0.0	2	4	US-08-361-024-3	Sequence 3, Appl1
20	0	0.0	2	4	US-08-361-024-3	Sequence 3, Appl1
21	0	0.0	2	4	US-09-472-035A-19	Sequence 19, Appl
22	0	0.0	2	4	US-09-472-035A-19	Sequence 19, Appl
23	0	0.0	2	4	US-09-472-035A-20	Sequence 20, Appl
24	0	0.0	2	4	US-09-472-035A-20	Sequence 20, Appl
25	0	0.0	2	4	US-09-411-862A-15	Sequence 15, Appl
26	0	0.0	2	4	US-09-411-862A-15	Sequence 15, Appl
27	0	0.0	2	4	US-09-411-862A-16	Sequence 16, Appl
28	0	0.0	2	4	US-09-411-862A-16	Sequence 16, Appl
29	0	0.0	2	4	US-09-227-782-16	Sequence 16, Appl
30	0	0.0	2	4	US-09-227-782-16	Sequence 16, Appl
31	0	0.0	2	5	PCT-US95-05758-16	Sequence 16, Appl
32	0	0.0	2	5	PCT-US95-05758-16	Sequence 16, Appl
33	0	0.0	3	1	US-07-791-213D-46	Sequence 46, Appl
34	0	0.0	3	1	US-07-791-213D-46	Sequence 46, Appl
35	0	0.0	3	1	US-07-791-213D-62	Sequence 62, Appl
36	0	0.0	3	1	US-07-791-213D-62	Sequence 62, Appl
37	0	0.0	3	1	US-08-268-679B-7	Sequence 7, Appl1
38	0	0.0	3	1	US-08-268-679B-7	Sequence 7, Appl1
39	0	0.0	3	1	US-08-295-743-22	Sequence 22, Appl
40	0	0.0	3	1	US-08-295-743-22	Sequence 22, Appl
41	0	0.0	3	1	US-08-602-036A-2	Sequence 2, Appl1
42	0	0.0	3	1	US-08-602-036A-2	Sequence 2, Appl1
43	0	0.0	3	1	US-08-293-150A-46	Sequence 46, Appl
44	0	0.0	3	1	US-08-293-150A-46	Sequence 46, Appl
45	0	0.0	3	1	US-08-293-150A-62	Sequence 62, Appl

#### ALIGNMENTS

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 11:26:42 ; Search time 27.3103 Seconds  
(without alignments)  
577.217 Million cell updates/sec

Title: US-09-877-606-8

Perfect score: 7

Sequence: 1 xxxxxxxx 7

Scoring table:

BLOSUM62	Xgapop 10.0 , Xgapext 0.5
	Ygapop 10.0 , Ygapext 0.5
	Fgapop 6.0 , Fgapext 7.0
	Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

```

-MODEL=frame+ p2n model -DEV=xlh
-O=Cygn2_1/USPRO.spool/US09877606/runat_18032003_134800_17645/app-query.fasta.1.1059
-DB=N_geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blomsurf2 -TRANS=human40.cdi
-LIST=45 -DOCCALIG=200 -THR_SCORE=Pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE-LOCAL -OUTFMT=plo -NORM-ext -HEAPSIZE=500 -MILEN=0 -MAXLEN=2000000000
-USER=US09877606 @CGN.1.1564 @runat.18032003.134800.17645 -NCPU=6 -ICPU=3
-NO_XLPEXT -NO_MMAP -LARGESURRY -NEG_SCORES=0 -NAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database :

```

1: N_geneseq_101002:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
25: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	0	0.0	1	20	AAK57131	Human mutant KCON3
2	0	0.0	1	20	AAK57131	Human mutant KCON3
3	0	0.0	1	24	ABK94535	Human BRCAL gene,
4	0	0.0	1	24	ABK94535	Human BRCAL gene,
5	0	0.0	2	24	ABK94446	Human BRCAL gene,
6	0	0.0	2	24	ABK94446	Human BRCAL gene,
7	0	0.0	2	24	ABK24295	Human microsatellite
8	0	0.0	2	24	ABK24295	Human microsatellite
9	0	0.0	3	9	AAK80743	Sequence encoding
10	0	0.0	3	16	AAK80743	Sequence encoding
11	0	0.0	3	16	AAK80743	Sequence encoding
12	0	0.0	3	16	AAK80743	Sequence encoding
13	0	0.0	3	17	AAK33326	Plasmid pEZ318.th
14	0	0.0	3	17	AAK33326	Plasmid pEZ318.th
15	0	0.0	3	17	AAK33326	CAPL trinucleotide
16	0	0.0	3	21	AAK94655	Human TUB gene pro
17	0	0.0	3	21	AAK94655	Human TUB gene pro
18	0	0.0	3	22	AAK20244	Human breast cancer
19	0	0.0	3	22	AAK20244	Human breast cancer
20	0	0.0	3	24	ABN73302	Bovine embryonic g
21	0	0.0	3	24	ABN73302	Bovine embryonic g
22	0	0.0	3	24	ABN73392	Bovine embryonic g
23	0	0.0	4	16	AAK81664	BGEF binding oligo
24	0	0.0	4	16	AAK81664	BGEF binding oligo
25	0	0.0	4	16	AAK81665	BGEF binding oligo
26	0	0.0	4	18	AAK6385	Probe for target n
27	0	0.0	4	18	AAK6385	Probe for target n
28	0	0.0	4	18	AAK6385	Immunostimulatory
29	0	0.0	4	18	AAK77252	Immunostimulatory
30	0	0.0	4	18	AAK77252	Immunostimulatory
31	0	0.0	4	22	AAK17191	Human breast cancer
32	0	0.0	4	22	AAK17191	Human breast cancer
33	0	0.0	4	22	AAK24357	Human breast cancer
34	0	0.0	4	22	AAK24357	Human breast cancer
35	0	0.0	4	22	AAK61450	Cyclin binding ham
36	0	0.0	4	22	AAK61450	Cyclin binding ham
37	0	0.0	4	24	ABK98116	Triple helix form
38	0	0.0	4	24	ABK98116	Triple helix form
39	0	0.0	4	24	ABK98117	Triple helix form
40	0	0.0	4	24	ABK98117	Triple helix form
41	0	0.0	4	24	ABK98122	Triple helix form
42	0	0.0	4	24	ABK98122	Triple helix form
43	0	0.0	4	24	ABK98123	Triple helix form
44	0	0.0	4	24	ABK98123	Triple helix form
45	0	0.0	4	24	ABK98163	Triple helix form

## ALIGNMENTS

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:35:13 ; Search time 326.517 Seconds  
(Without alignments) 623.917 Million cell updates/sec

Title: US-09-877-606-8  
Perfect score: 7  
Sequence: 1 xxxxxxx 7

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xjh  
-O=/cgs2\_1/uspro.spool/US09877606/runat\_134800\_17655/app\_query.fasta.1.1059  
-DB=GenBml -OPMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USBR=US09877606.@CGN.1.1.3852.@runat.18032003.134800.17655 -NCPU=6 -ICPU=3  
-NO\_XLIPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenBml:\*  
1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vl:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.in:\*  
18: em.hum:\*  
19: em.mu:\*  
20: em.om:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.pl:\*  
25: em.ph:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vl:\*  
30: em.htg.hum:\*  
31: em.htg.in:\*  
32: em.htg.other:\*  
33: em.htg.mus:\*  
34: em.htg.pln:\*  
35: em.htg.rtd:\*  
36: em.htg.mam:\*  
37: em.htg.vrt:\*  
38: em.sy:\*  
39: em.htgo.hum:\*  
40: em.htgo.mus:\*  
41: em.htgo.other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	0	0.0	1	6	A95992 Sequence 25
2	0	0.0	1	6	A95992 Sequence 25
3	0	0.0	1	6	A96002 Sequence 35
4	0	0.0	1	6	A96002 Sequence 35
5	0	0.0	1	6	A96012 Sequence 45
6	0	0.0	1	6	A96012 Sequence 45
7	0	0.0	1	6	A96022 Sequence 55
8	0	0.0	1	6	A96022 Sequence 55
9	0	0.0	1	6	A96040 Sequence 73
10	0	0.0	1	6	A96040 Sequence 73
11	0	0.0	1	6	A96058 Sequence 91
12	0	0.0	1	6	A96058 Sequence 91
13	0	0.0	1	6	A96076 Sequence 10
14	0	0.0	1	6	A96076 Sequence 10
15	0	0.0	1	6	A96086 Sequence 11
16	0	0.0	1	6	A96086 Sequence 11
17	0	0.0	1	6	A96096 Sequence 12
18	0	0.0	1	6	A96096 Sequence 12
19	0	0.0	1	6	A96112 Sequence 14
20	0	0.0	1	6	A96112 Sequence 14
21	0	0.0	1	6	A96136 Sequence 16
22	0	0.0	1	6	A96136 Sequence 16
23	0	0.0	1	6	A96160 Sequence 19
24	0	0.0	1	6	A96160 Sequence 19
25	0	0.0	1	6	A96170 Sequence 20
26	0	0.0	1	6	A96170 Sequence 20
27	0	0.0	1	6	A96190 Sequence 22
28	0	0.0	1	6	A96190 Sequence 22
29	0	0.0	1	6	A96196 Sequence 22
30	0	0.0	1	6	A96196 Sequence 22
31	0	0.0	1	6	A96204 Sequence 23
32	0	0.0	1	6	A96204 Sequence 23
33	0	0.0	1	6	A96222 Sequence 25
34	0	0.0	1	6	A96222 Sequence 25
35	0	0.0	1	6	A96240 Sequence 27
36	0	0.0	1	6	A96240 Sequence 27
37	0	0.0	1	6	A96258 Sequence 29
38	0	0.0	1	6	A96258 Sequence 29
39	0	0.0	1	6	A96300 Sequence 33
40	0	0.0	1	6	A96300 Sequence 33
41	0	0.0	1	6	A96318 Sequence 35
42	0	0.0	1	6	A96318 Sequence 35
43	0	0.0	1	6	A96328 Sequence 36
44	0	0.0	1	6	A96328 Sequence 36
45	0	0.0	1	6	A96354 Sequence 38

## ALIGNMENTS

SEQ ID NO 46  
LENGTH: 1  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-46

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 K 1

RESULT 12  
US-09-982-172-69  
Sequence 69, Application US/09982172  
Patent No. US20020137119A1  
GENERAL INFORMATION:  
APPLICANT: Emil Israel Katz  
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
FILE REFERENCE: 01/22283  
CURRENT APPLICATION NUMBER: US/09/982,172  
CURRENT FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 69  
LENGTH: 1  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-69

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 K 1

RESULT 13  
US-09-982-172-80  
Sequence 80, Application US/09982172  
Patent No. US20020137119A1  
GENERAL INFORMATION:  
APPLICANT: Emil Israel Katz  
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
FILE REFERENCE: 01/22283  
CURRENT APPLICATION NUMBER: US/09/982,172  
CURRENT FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 80  
LENGTH: 1  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-80

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 K 1

RESULT 14  
US-09-982-172-81  
Sequence 81, Application US/09982172  
Patent No. US20020137119A1  
GENERAL INFORMATION:  
APPLICANT: Emil Israel Katz  
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
FILE REFERENCE: 01/22283  
CURRENT APPLICATION NUMBER: US/09/982,172  
CURRENT FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 81  
LENGTH: 1  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-81

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 K 1

RESULT 15  
US-09-982-172-83  
Sequence 83, Application US/09982172  
Patent No. US20020137119A1  
GENERAL INFORMATION:  
APPLICANT: Emil Israel Katz  
TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIES  
TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERATING  
FILE REFERENCE: 01/22283  
CURRENT APPLICATION NUMBER: US/09/982,172  
CURRENT FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 253  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 83  
LENGTH: 1  
TYPE: PRT  
ORGANISM: Artificial sequence  
FEATURE:  
OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-83

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 K 1

Search completed: March 18, 2003, 14:52:58  
Job time : 2.62069 secs

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 K 1

RESULT 7  
US-09-982-172-19  
; Sequence 19, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE  
; TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-19

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 K 1

RESULT 8  
US-09-982-172-31  
; Sequence 31, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE  
; TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-31

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 K 1

RESULT 9  
US-09-982-172-35  
; Sequence 35, Application US/09982172

Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBO  
; TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 35  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-35

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 R 1

RESULT 10  
US-09-982-172-37  
; Sequence 37, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBO  
; TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 37  
; LENGTH: 1  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Computer generated synthetic peptide  
US-09-982-172-37

Query Match 0.0%; Score 0; DB 10; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 K 1

RESULT 11  
US-09-982-172-46  
; Sequence 46, Application US/09982172  
; Patent No. US20020137119A1  
; GENERAL INFORMATION:  
; APPLICANT: Emil Israel Katz  
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBO  
; TITLE OF INVENTION: DIRECTED THERAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE  
; FILE REFERENCE: 01/22283  
; CURRENT APPLICATION NUMBER: US/09/982,172  
; CURRENT FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 253  
; SOFTWARE: PatentIn version 3.1



```
; TITLE OF INVENTION: Stimulation Of Cartilage Growth With Agonists
; TITLE OF INVENTION: Of The No. US20020042373A1-Proteolytically Activated Thrombin Re
; FILE REFERENCE: 3033.1003-001
; CURRENT APPLICATION NUMBER: US/09/909,348
; PRIOR FILING DATE: 2001-07-19
; PRIOR APPLICATION NUMBER: US 60/219,800
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 14
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide fragment of Thrombin
; NAME/KEY: VARIANT
; LOCATION: (1)...(14)
; OTHER INFORMATION: xaa at position six is Glu or Gln
; OTHER INFORMATION: xaa at position thirteen is Phe, Met, Leu, His or
US-09-909-348-4
```

```
Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 X 1
DB      1 V 1
```

```
RESULT 3
US-09-982-172-3
; Sequence 3, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-3
```

```
Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 X 1
DB      1 K 1
```

```
RESULT 4
US-09-982-172-4
; Sequence 4, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBODIE
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENERAT
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
```

```
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-4
```

```
Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 X 1
DB      1 K 1
```

```
RESULT 5
US-09-982-172-9
; Sequence 9, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBO
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-9
```

```
Query Match          0.0%; Score 0; DB 10; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY      1 X 1
DB      1 K 1
```

```
RESULT 6
US-09-982-172-11
; Sequence 11, Application US/09982172
; Patent No. US20020137119A1
; GENERAL INFORMATION:
; APPLICANT: Emil Israel Katz
; TITLE OF INVENTION: PEPTIDES REPRESENTATIVE OF POLYPEPTIDES OF INTEREST AND ANTIBO
; TITLE OF INVENTION: DIRECTED THEREAGAINST, AND METHODS, SYSTEMS AND KITS FOR GENE
; FILE REFERENCE: 01/22283
; CURRENT APPLICATION NUMBER: US/09/982,172
; CURRENT FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 1
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Computer generated synthetic peptide
US-09-982-172-11
```

```
Query Match          0.0%; Score 0; DB 10; Length 1;
```

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:36:16 ; Search time 1.62069 Seconds  
(without alignments)  
199.079 Million cell updates/sec

Title: US-09-877-606-8  
Perfect score: 7  
Sequence: 1 XXXXXXX 7

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues  
Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:\*  
1: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB pep:\*  
2: /cgn2\_6/ptodata/2/pubppaa/PC7\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB pep:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB pep:\*  
7: /cgn2\_6/ptodata/2/pubppaa/PC7US\_PUBCOMB pep:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB pep:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB pep:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB pep:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB pep:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB pep:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB pep:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	9 US-10-024-860-4	Sequence 4, Appl1
2	0	0.0	1	US-09-909-348-4	Sequence 4, Appl1
3	0	0.0	1	US-09-982-172-3	Sequence 3, Appl1
4	0	0.0	1	US-09-982-172-4	Sequence 9, Appl1
5	0	0.0	1	US-09-982-172-9	Sequence 11, Appl1
6	0	0.0	1	US-09-982-172-11	Sequence 19, Appl1
7	0	0.0	1	US-09-982-172-19	Sequence 31, Appl1
8	0	0.0	1	US-09-982-172-31	Sequence 35, Appl1
9	0	0.0	1	US-09-982-172-35	Sequence 37, Appl1
10	0	0.0	1	US-09-982-172-37	Sequence 46, Appl1
11	0	0.0	1	US-09-982-172-46	Sequence 69, Appl1
12	0	0.0	1	US-09-982-172-69	Sequence 80, Appl1
13	0	0.0	1	US-09-982-172-80	Sequence 81, Appl1
14	0	0.0	1	US-09-982-172-81	Sequence 83, Appl1
15	0	0.0	1	US-09-982-172-83	Sequence 86, Appl1
16	0	0.0	1	US-09-982-172-86	Sequence 93, Appl1
17	0	0.0	1	US-09-982-172-93	Sequence 95, Appl1
18	0	0.0	1	US-09-982-172-95	Sequence 106, Appl1
19	0	0.0	1	US-09-982-172-106	

20	0	0.0	1	10 US-09-982-172-112	Sequence 112, App
21	0	0.0	1	10 US-09-982-172-120	Sequence 120, App
22	0	0.0	1	10 US-09-982-172-126	Sequence 126, App
23	0	0.0	1	10 US-09-982-172-148	Sequence 148, App
24	0	0.0	1	10 US-09-982-172-149	Sequence 149, App
25	0	0.0	1	10 US-09-982-172-155	Sequence 155, App
26	0	0.0	1	10 US-09-982-172-160	Sequence 160, App
27	0	0.0	1	10 US-09-982-172-172	Sequence 172, App
28	0	0.0	1	10 US-09-982-172-173	Sequence 173, App
29	0	0.0	1	10 US-09-982-172-175	Sequence 175, App
30	0	0.0	1	10 US-09-982-172-189	Sequence 189, App
31	0	0.0	1	10 US-09-982-172-190	Sequence 190, App
32	0	0.0	1	10 US-09-982-172-191	Sequence 191, App
33	0	0.0	1	10 US-09-982-172-195	Sequence 195, App
34	0	0.0	1	10 US-09-982-172-200	Sequence 200, App
35	0	0.0	1	10 US-09-982-172-211	Sequence 211, App
36	0	0.0	2	9 US-09-554-000-52	Sequence 52, Appl1
37	0	0.0	2	9 US-09-887-853-9	Sequence 9, Appl1
38	0	0.0	2	9 US-10-061-395-70	Sequence 70, Appl1
39	0	0.0	2	9 US-10-005-438-3	Sequence 3, Appl1
40	0	0.0	2	9 US-10-119-417-16	Sequence 16, Appl1
41	0	0.0	2	10 US-09-816-737-14	Sequence 14, Appl1
42	0	0.0	2	10 US-09-821-883-14	Sequence 14, Appl1
43	0	0.0	2	10 US-09-867-847-33	Sequence 33, Appl1
44	0	0.0	2	10 US-09-867-847-34	Sequence 34, Appl1
45	0	0.0	2	10 US-09-867-847-35	Sequence 35, Appl1

## ALIGNMENTS

RESULT 1  
US-10-024-860-4  
Sequence 4, Application US/10024860  
Patent No. US20020172692A1  
GENERAL INFORMATION:  
APPLICANT: Cohen, Joseph  
TITLE OF INVENTION: Vaccine Composition Against Malaria  
FILE REFERENCE: B45088C2  
CURRENT APPLICATION NUMBER: US/10/024, 860  
CURRENT FILING DATE: 2001-12-18  
PRIOR APPLICATION NUMBER: 09/826, 513  
PRIOR FILING DATE: 2001-04-05  
PRIOR APPLICATION NUMBER: 09/230, 629  
PRIOR FILING DATE: 1999-01-26  
PRIOR APPLICATION NUMBER: GB 9616351.4  
PRIOR FILING DATE: 1996-08-02  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 4  
LENGTH: 1  
TYPE: PRT  
ORGANISM: Hepatitis B  
US-10-024-860-4

Query Match 0.0%; Score 0; DB 9; Length 1;  
Best Local Similarity 0.0%; Pred. NO. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0;  
QY 1 X 1  
Db 1 R 1

RESULT 2  
US-09-909-348-4  
Sequence 4, Application US/09909348  
Patent No. US20020042373A1  
GENERAL INFORMATION:  
APPLICANT: Carney, Darrell H.  
APPLICANT: Crowley, Roger S.  
APPLICANT: Stierberg, Janet  
APPLICANT: Bergmann, John

**THIS PAGE BLANK (USPTO)**

REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0009.31  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT  
US-08-496-847-23

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 C 1

RESULT 15  
US-08-496-847-25  
Sequence 25, Application US/08496847  
Patent No. 5795864  
GENERAL INFORMATION:  
APPLICANT: Amstutz, Gary A.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Gohil, Kishorchandra  
APPLICANT: Adriaenssens, Peter I.  
APPLICANT: Kristipati, Ramasharma  
TITLE OF INVENTION: METHODS AND  
TITLE OF INVENTION: FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Dehlinger & Associates  
STREET: 350 Cambridge Avenue, Suite 250  
CITY: Palo Alto  
STATE: CA  
COUNTRY: US  
ZIP: 94306-1546  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,847  
FILING DATE: 27-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A  
REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0009.31  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-324-0880  
TELEFAX: 650-324-0960  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT  
US-08-496-847-25

Query Match 0.0%; Score 0; DB 1; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 R 1

Search completed: March 18, 2003, 14:50:06  
Job time : 2.68966 secs

Patent No. 5792629  
GENERAL INFORMATION:  
APPLICANT: MORISHITA, Hideaki  
APPLICANT: KANAMORI, Toshiro  
APPLICANT: NOBUHARA, Masahiro  
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
TITLE OF INVENTION: TREATING USING THE SAME  
NUMBER OF SEQUENCES: 110  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/293,150A  
FILING DATE: 19-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/791,213  
FILING DATE: 13-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-306745  
FILING DATE: 13-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Meuth, Donna M.  
REGISTRATION NUMBER: 36,607  
REFERENCE/DOCKET NUMBER: 029650-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-293-150A-24

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 x 1  
DB 1 c 1

RESULT 13  
US-08-293-150A-40  
Sequence 40, Application US/08293150A  
Patent No. 5792629  
GENERAL INFORMATION:  
APPLICANT: MORISHITA, Hideaki  
APPLICANT: KANAMORI, Toshiro  
APPLICANT: NOBUHARA, Masahiro  
TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
TITLE OF INVENTION: TREATING USING THE SAME  
NUMBER OF SEQUENCES: 110  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/293,150A  
FILING DATE: 19-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/791,213  
FILING DATE: 13-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-306745  
FILING DATE: 13-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Meuth, Donna M.  
REGISTRATION NUMBER: 36,607  
REFERENCE/DOCKET NUMBER: 029650-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-293-150A-40

STATE: Virginia  
COUNTRY: United States  
ZIP: 2213-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/293,150A  
FILING DATE: 19-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/791,213  
FILING DATE: 13-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-306745  
FILING DATE: 13-NOV-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Meuth, Donna M.  
REGISTRATION NUMBER: 36,607  
REFERENCE/DOCKET NUMBER: 029650-049  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-293-150A-40

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 x 1  
DB 1 c 1

RESULT 14  
US-08-496-847-23  
Sequence 23, Application US/08496847  
Patent No. 5795864  
GENERAL INFORMATION:  
APPLICANT: Amstutz, Gary A.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Gohl, Kishorchandra  
APPLICANT: Adriaenssens, Peter I.  
APPLICANT: Kristipati, Ramasharma  
TITLE OF INVENTION: METHODS AND  
TITLE OF INVENTION: FORMULATIONS FOR PREVENTING PROGRESSION OF NEUROPATHIC PAIN  
NUMBER OF SEQUENCES: 36  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/496,847  
FILING DATE: 27-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/433.037  
FILING DATE: 03-MAY-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9108Z  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 12:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-433-037-12

Query Match  
Best Local Similarity 0.0%; Score 0; DB 1; Length 1;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 x 1  
Db 1 L 1

RESULT 10  
US-08-448-606-4  
Sequence 4, Application US/08448606  
Patent No. 5721114  
GENERAL INFORMATION:  
APPLICANT: Abrahams n, Lars  
APPLICANT: Holmgren, Erik  
APPLICANT: Kalder n, Christlma  
APPLICANT: Lake, Mats  
APPLICANT: Mikaelsson, sa  
APPLICANT: Sejlitz, Torsten  
TITLE OF INVENTION: Expression System For Producing  
TITLE OF INVENTION: Apolipoprotein AI-M  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pollock, Vande Sande & Priddy  
STREET: 1990 M Street, N.W., Suite 800  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448.606  
FILING DATE: 25-AUG-1995  
CLASSIFICATION: 436  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/SE93/01061  
FILING DATE: 09-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: SE 9203753-0  
FILING DATE: 11-DEC-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Amernick, Burton A.  
REGISTRATION NUMBER: 24,852  
REFERENCE/DOCKET NUMBER: 0151/00121  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)331-7111  
TELEFAX: (202) 293-6229  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acid  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-448-606-4

Query Match  
Best Local Similarity 0.0%; Score 0; DB 1; Length 1;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 x 1  
Db 1 Q 1

RESULT 11  
US-07-869-933-16  
Sequence 16, Application US/07869933  
Patent No. 5770396  
GENERAL INFORMATION:  
APPLICANT: Kinet, Jean-Pierre  
TITLE OF INVENTION: ISOLATION, CHARACTERIZATION, AND USE OF  
TITLE OF INVENTION: THE HUMAN B SUBUNIT OF THE HIGH AFFINITY RECEPTOR FOR  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 1800 Diagonal Road, Suite 500  
CITY: Alexandria  
STATE: VA  
COUNTRY: USA  
ZIP: 22313-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/869.933  
FILING DATE: 19920416  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Benti, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 40399/154 NTHD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703)836-9300  
TELEFAX: (703)683-4109  
TELEX: 899149  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-869-933-16

Query Match  
Best Local Similarity 0.0%; Score 0; DB 1; Length 1;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 x 1  
Db 1 M 1

RESULT 12  
US-08-293-150A-24  
Sequence 24, Application US/08293150A

Db 1 R 1

RESULT 7  
US-08-049-794-23  
; Sequence 23, Application US/08049794  
; Patent No. 5587454  
; GENERAL INFORMATION:  
; APPLICANT: JUSTICE, ALAN  
; APPLICANT: SINGH, TEJINDER  
; APPLICANT: GOHIL, KISHOR C  
; APPLICANT: VALENTINO, KAREN L  
; APPLICANT: MILJANICH, GEORGE P  
; TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND  
; TITLE OF INVENTION: ENHANCING OPIATE ANALGESIA  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Offices of Peter Dehlinger,  
; STREET: 350 Cambridge Avenue, Suite 300  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,794  
; FILING DATE: 19930415  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/814,759  
; FILING DATE: 30-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stratford, Carol A.  
; REGISTRATION NUMBER: 34,444  
; REFERENCE/DOCKET NUMBER: 5865-0009.30  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 23:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE  
; INDIVIDUAL ISOLATE: 32  
US-08-049-794-23

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 C 1

RESULT 8  
US-08-049-794-25  
; Sequence 25, Application US/08049794  
; Patent No. 5587454  
; GENERAL INFORMATION:  
; APPLICANT: JUSTICE, ALAN  
; APPLICANT: SINGH, TEJINDER  
; APPLICANT: GOHIL, KISHOR C  
; APPLICANT: VALENTINO, KAREN L

APPLICANT: MILJANICH, GEORGE P  
; TITLE OF INVENTION: METHODS OF PRODUCING ANALGESIA AND  
; TITLE OF INVENTION: ENHANCING OPIATE ANALGESIA  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Offices of Peter Dehlinger  
; STREET: 350 Cambridge Avenue, Suite 300  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/049,794  
; FILING DATE: 19930415  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/814,759  
; FILING DATE: 30-DEC-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Stratford, Carol A.  
; REGISTRATION NUMBER: 34,444  
; REFERENCE/DOCKET NUMBER: 5865-0009.30  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 25:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: CONOPEPTIDE GROUP 1 FRAGMENT, PAGE  
; INDIVIDUAL ISOLATE: 32  
US-08-049-794-25

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 R 1

RESULT 9  
US-08-433-037-12  
; Sequence 12, Application US/08433037  
; Patent No. 5707828  
; GENERAL INFORMATION:  
; APPLICANT: Sreekrishna, Kotikanyadan  
; APPLICANT: Barr, Kathryn A.  
; APPLICANT: Brierley, Russell A.  
; APPLICANT: Thill, Gregory P.  
; APPLICANT: Tschopp, Juerg F.  
; TITLE OF INVENTION: EXPRESSION OF HUMAN SERUM ALBUMIN IN  
; TITLE OF INVENTION: PICHIA PASTORIS  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Scully, Scott, Murphy & Presser  
; STREET: 400 Garden City Plaza  
; CITY: Garden City  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 11530-0299  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acid  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: modified site  
LOCATION:  
IDENTIFICATION METHOD:  
OTHER INFORMATION: /note = "Xaa is modified amino acid as  
US-08-174-365A-57  
OTHER INFORMATION: described in specification"  
Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 X 1  
DB 1 X 1  
RESULT 5  
US-07-789-913-23  
Sequence 23, Application US/07789913  
Patent No. 5559095  
GENERAL INFORMATION:  
APPLICANT: Miljanich, George P.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Fox, James A.  
APPLICANT: Valentino, Karen L.  
APPLICANT: Bitner, Robert S.  
APPLICANT: Yamashiro, Donald H.  
TITLE OF INVENTION: Delayed Treatment Method of Reducing  
TITLE OF INVENTION: Ischemia-Related Neuronal Damage  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Peter Dehlinger  
STREET: 350 Cambridge Avenue, Suite 300  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/789,913  
FILING DATE: 19911112  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/561,766  
FILING DATE: 02-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/440,094  
FILING DATE: 22-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A.  
REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0005.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0960  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: peptide fragment used in the claims  
US-07-789-913-23  
Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
DB 1 C 1  
RESULT 6  
US-07-789-913-25  
Sequence 25, Application US/07789913  
Patent No. 5559095  
GENERAL INFORMATION:  
APPLICANT: Miljanich, George P.  
APPLICANT: Bowersox, Stephen S.  
APPLICANT: Fox, James A.  
APPLICANT: Valentino, Karen L.  
APPLICANT: Bitner, Robert S.  
APPLICANT: Yamashiro, Donald H.  
TITLE OF INVENTION: Delayed Treatment Method of Reducing  
TITLE OF INVENTION: Ischemia-Related Neuronal Damage  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Law Offices of Peter Dehlinger  
STREET: 350 Cambridge Avenue, Suite 300  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/789,913  
FILING DATE: 19911112  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/561,766  
FILING DATE: 02-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/440,094  
FILING DATE: 22-NOV-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Stratford, Carol A.  
REGISTRATION NUMBER: 34,444  
REFERENCE/DOCKET NUMBER: 5865-0005.30  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 324-0960  
TELEFAX: (415) 324-0960  
INFORMATION FOR SEQ ID NO: 25:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: both  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
INDIVIDUAL ISOLATE: peptide fragment used in the claims  
US-07-789-913-25  
Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1



; Sequence 24, Application US/07791213D  
; Patent No. 5409895  
; GENERAL INFORMATION:  
; APPLICANT: MORISHITA, Hideaki  
; APPLICANT: KANAMORI, Toshiro  
; APPLICANT: NOBUHARA, Masahiro  
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States  
; ZIP: 22133-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/791,213D  
; FILING DATE: 13-NOV-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-306745  
; FILING DATE: 13-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Neuth, Donna M  
; REGISTRATION NUMBER: 36,607  
; REFERENCE/DOCKET NUMBER: 029650-032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 24:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-791-213D-24

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 C 1

RESULT 3  
US-07-791-213D-40  
; Sequence 40, Application US/07791213D  
; Patent No. 5409895  
; GENERAL INFORMATION:  
; APPLICANT: MORISHITA, Hideaki  
; APPLICANT: KANAMORI, Toshiro  
; APPLICANT: NOBUHARA, Masahiro  
; TITLE OF INVENTION: POLYPEPTIDE, DNA FRAGMENT ENCODING THE  
; TITLE OF INVENTION: SAME AND PROCESS FOR PRODUCING THE SAME, AND ENZYME  
; TITLE OF INVENTION: INHIBITION PROCESS, DRUG COMPOSITION AND METHODS OF  
; NUMBER OF SEQUENCES: 108  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Burns, Doane, Swecker & Mathis  
; STREET: P.O. Box 1404  
; CITY: Alexandria  
; STATE: Virginia  
; COUNTRY: United States

; ZIP: 22133-1404  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/791,213D  
; FILING DATE: 13-NOV-1991  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 2-306745  
; FILING DATE: 13-NOV-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Neuth, Donna M  
; REGISTRATION NUMBER: 36,607  
; REFERENCE/DOCKET NUMBER: 029650-032  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 836-6620  
; TELEFAX: (703) 836-2021  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-07-791-213D-40

Query Match 0.0%; Score 0; DB 1; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
DB 1 C 1

RESULT 4  
US-08-174-365A-57  
; Sequence 57, Application US/08174365A  
; Patent No. 5478809  
; GENERAL INFORMATION:  
; APPLICANT: Seichi TANIDA et al.  
; TITLE OF INVENTION: 4-THIAHEPTANOIC ACID DERIVATIVES  
; NUMBER OF SEQUENCES: 106  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Wenderoth, Lind & Ponack  
; STREET: 805 Fifteenth Street, N.W., #700  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.5 inch, 144 mb  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: MS-DOS  
; SOFTWARE: Wordperfect 5.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/174,365A  
; FILING DATE: December 28, 1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION NUMBER:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Warren M. Cheek, Jr.  
; REGISTRATION NUMBER: 33,367  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-371-8850  
; TELEFAX:  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 57:

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:33:36 ; Search time 1.68966 Seconds  
(without alignments)  
121.895 Million cell updates/sec

Title: US-09-877-606-8  
Perfect score: 7  
Sequence: 1 xxxxxxxx 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Issued Patents\_AA:\*  
2: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfill1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0.0	0.0	1	US-07-820-154A-12	Sequence 12, Appl
2	0.0	0.0	1	US-07-791-213D-24	Sequence 24, Appl
3	0.0	0.0	1	US-07-791-213D-40	Sequence 40, Appl
4	0.0	0.0	1	US-08-174-365A-57	Sequence 57, Appl
5	0.0	0.0	1	US-07-789-913-23	Sequence 23, Appl
6	0.0	0.0	1	US-07-789-913-25	Sequence 25, Appl
7	0.0	0.0	1	US-08-049-794-23	Sequence 23, Appl
8	0.0	0.0	1	US-08-049-794-25	Sequence 25, Appl
9	0.0	0.0	1	US-08-433-037-12	Sequence 12, Appl
10	0.0	0.0	1	US-08-448-606-4	Sequence 4, Appl
11	0.0	0.0	1	US-07-869-933-16	Sequence 16, Appl
12	0.0	0.0	1	US-08-293-150A-24	Sequence 24, Appl
13	0.0	0.0	1	US-08-293-150A-40	Sequence 40, Appl
14	0.0	0.0	1	US-08-496-847-23	Sequence 23, Appl
15	0.0	0.0	1	US-08-496-847-25	Sequence 25, Appl
16	0.0	0.0	1	US-08-742-774-23	Sequence 23, Appl
17	0.0	0.0	1	US-08-742-774-25	Sequence 25, Appl
18	0.0	0.0	1	US-08-675-354-23	Sequence 23, Appl
19	0.0	0.0	1	US-08-675-354-25	Sequence 25, Appl
20	0.0	0.0	1	US-08-097-554A-12	Sequence 12, Appl
21	0.0	0.0	1	US-08-965-918-23	Sequence 23, Appl
22	0.0	0.0	1	US-08-965-918-25	Sequence 25, Appl
23	0.0	0.0	1	US-09-138-439-23	Sequence 23, Appl
24	0.0	0.0	1	US-09-138-439-25	Sequence 25, Appl
25	0.0	0.0	1	US-08-480-640A-12	Sequence 12, Appl
26	0.0	0.0	1	US-08-613-400A-23	Sequence 23, Appl
27	0.0	0.0	1	US-08-613-400A-25	Sequence 25, Appl

28	0.0	0.0	1	US-08-801-092-10	Sequence 10, Appl
29	0.0	0.0	1	US-08-801-092-17	Sequence 17, Appl
30	0.0	0.0	1	US-08-801-092-24	Sequence 24, Appl
31	0.0	0.0	1	US-08-801-092-31	Sequence 31, Appl
32	0.0	0.0	1	US-08-801-092-38	Sequence 38, Appl
33	0.0	0.0	1	US-08-801-092-45	Sequence 45, Appl
34	0.0	0.0	1	US-09-298-017-23	Sequence 23, Appl
35	0.0	0.0	1	US-09-298-017-25	Sequence 25, Appl
36	0.0	0.0	1	US-08-295-802-12	Sequence 12, Appl
37	0.0	0.0	1	US-09-392-979A-23	Sequence 23, Appl
38	0.0	0.0	1	US-09-392-979A-25	Sequence 25, Appl
39	0.0	0.0	1	US-09-103-663-16	Sequence 16, Appl
40	0.0	0.0	1	US-08-488-237A-12	Sequence 12, Appl
41	0.0	0.0	1	US-09-117-927-5	Sequence 5, Appl
42	0.0	0.0	1	US-08-375-992A-12	Sequence 12, Appl
43	0.0	0.0	1	US-09-315-113-10	Sequence 10, Appl
44	0.0	0.0	1	US-09-315-113-17	Sequence 17, Appl
45	0.0	0.0	1	US-09-315-113-24	Sequence 24, Appl

## ALIGNMENTS

```

RESULT 1
US-07-820-154A-12
; Sequence 12, Application US/07820154A
; Patent No. 5382425
;
GENERAL INFORMATION:
; APPLICANT: Cochran Ph.D., Mark D
; APPLICANT: Junker M.S., David E
; TITLE OF INVENTION: Recombinant Swinepox Virus
; NUMBER OF SEQUENCES: 40
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: John P. White
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10112
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/820,154A
; FILING DATE: 19920113
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)977-9550
; TELEFAX: (212)664-0525
; TELEX: 422523
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
Query Match 0.0%; Score 0; DB 1; Length 1;
Best Local Similarity 0.0%; Pred. No. 0;
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 X 1
Db 1 V 1
RESULT 2
US-07-791-213D-24

```

XX polymorphic nucleic acid sequences, useful in genetic testing and  
PT therapy -  
XX  
PS Claim 29; Page 2641; 2653pp; English.  
XX  
CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide  
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
CC AAM53114 to AAM53329 represent peptides related to human polymorphic  
CC polynucleotide sequences. The sequences can be used in gene and protein  
CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of polymorphic polypeptides.  
CC For example, (I) may be used to treat disorders by rectifying mutations  
CC or deletions in a patient's genome that affect the activity of  
CC polypeptides by expressing inactive proteins or to supplement the  
CC patients own production of polypeptide. Additionally, (I) and its  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids  
CC in samples, and therefore which patients may be in need of restorative  
CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
CC production of antibodies specific for polymorphic polypeptides. The  
CC antibodies may also be used to down regulate expression and activity.  
CC The antibodies may also be used as diagnostic agents for detecting the  
CC presence of polymorphic polypeptides in samples.  
XX  
SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 L 1

RESULT 15  
AAM53328

ID AAM53328 standard; Peptide; 1 AA.

AC AAM53328;

DT 09-NOV-2001 (first entry)

DE Human nonconservative amino acid changing SNP related peptide SEQ:7023.

KM Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
KM protein therapy; vaccine; probe; diagnostic assay; detection;  
KM quantitation; restorative therapy; polymorphic.

OS Homo sapiens.

PN WO200140521-A2.

PD 07-JUN-2001.

PF 30-NOV-2000; 2000WO-US32758.

PR 30-NOV-1999; 99US-0168138.

PR 29-NOV-2000; 2000US-0726173.

PA (CURA-) CURAGEN CORP.

PI Shimkets RA, Leach M;

DR WPI; 2001-356160/37.

PT Polymorphic nucleic acid sequences, useful in genetic testing and  
therapy -

PS Claim 29; Page 2652; 2653pp; English.  
XX

CC AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide  
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
CC AAM53114 to AAM53329 represent peptides related to human polymorphic  
CC polynucleotide sequences. The sequences can be used in gene and protein  
CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of polymorphic polypeptides.  
CC For example, (I) may be used to treat disorders by rectifying mutations  
CC or deletions in a patient's genome that affect the activity of  
CC polypeptides by expressing inactive proteins or to supplement the  
CC patients own production of polypeptide. Additionally, (I) and its  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids  
CC in samples, and therefore which patients may be in need of restorative  
CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
CC production of antibodies specific for polymorphic polypeptides. The  
CC antibodies may also be used to down regulate expression and activity.  
CC The antibodies may also be used as diagnostic agents for detecting the  
CC presence of polymorphic polypeptides in samples.  
XX  
SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 V 1

Search completed: March 18, 2003, 14:33:28  
Job time : 7 secs

PD 07-JUN-2001.  
XX  
PF 30-NOV-2000; 2000WO-US32758.  
XX  
PR 30-NOV-1999; 99US-0168138.  
PR 29-NOV-2000; 2000US-0726173.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shimkets RA, Leach M;  
XX  
DR WPI; 2001-356160/37.  
PT Polymorphic nucleic acid sequences, useful in genetic testing and  
PR therapy -  
XX  
PS Claim 29; Page 2619; 2653pp; English.  
XX  
CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide  
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
CC AA53114 to AA53329 represent peptides related to human polymorphic  
CC polynucleotide sequences. The sequences can be used in gene and protein  
CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of polymorphic polypeptides.  
CC For example, (I) may be used to treat disorders by rectifying mutations  
CC or deletions in a patient's genome that affect the activity of  
CC polypeptides by expressing inactive proteins or to supplement the  
CC patient's own production of polypeptide. Additionally, (I) and its  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids  
CC in samples, and therefore which patients may be in need of restorative  
CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
CC production of antibodies specific for polymorphic polypeptides. The  
CC antibodies may also be used to down regulate expression and activity.  
CC The antibodies may also be used as diagnostic agents for detecting the  
CC presence of polymorphic polypeptides in samples.  
XX  
SQ Sequence 1 AA;  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
DB 1 H 1  
RESULT 13  
AA53290  
ID AA53290 standard; Peptide: 1 AA.  
XX  
AC AA53290;  
XX  
DT 09-NOV-2001 (first entry)  
XX  
DE Human nonconservative amino acid changing SNP related peptide SEQ:6985.  
XX  
KW Human: single nucleotide polymorphism; SNP; genome; gene therapy;  
KW protein therapy; vaccine; probe; diagnostic assay; detection;  
KW quantitation; restorative therapy; polymorphic.  
XX  
OS Homo sapiens.  
XX  
PN WO200140521-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 30-NOV-2000; 2000WO-US32758.  
XX  
PR 30-NOV-1999; 99US-0168138.  
PR 29-NOV-2000; 2000US-0726173.

XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shimkets RA, Leach M;  
XX  
DR WPI; 2001-356160/37.  
XX  
PT Polymorphic nucleic acid sequences, useful in genetic testing and  
PR therapy -  
XX  
PS Claim 29; Page 2641; 2653pp; English.  
XX  
CC AA173060 to AA179867 represent isolated human polymorphic polynucleotide  
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
CC AA53114 to AA53329 represent peptides related to human polymorphic  
CC polynucleotide sequences. The sequences can be used in gene and protein  
CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of polymorphic polypeptides.  
CC For example, (I) may be used to treat disorders by rectifying mutations  
CC or deletions in a patient's genome that affect the activity of  
CC polypeptides by expressing inactive proteins or to supplement the  
CC patient's own production of polypeptide. Additionally, (I) and its  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids  
CC in samples, and therefore which patients may be in need of restorative  
CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
CC production of antibodies specific for polymorphic polypeptides. The  
CC antibodies may also be used to down regulate expression and activity.  
CC The antibodies may also be used as diagnostic agents for detecting the  
CC presence of polymorphic polypeptides in samples.  
XX  
SQ Sequence 1 AA;  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
DB 1 L 1  
RESULT 14  
AA53291  
ID AA53291 standard; Peptide: 1 AA.  
XX  
AC AA53291;  
XX  
DT 09-NOV-2001 (first entry)  
XX  
DE Human nonconservative amino acid changing SNP related peptide SEQ:6986.  
XX  
KW Human: single nucleotide polymorphism; SNP; genome; gene therapy;  
KW protein therapy; vaccine; probe; diagnostic assay; detection;  
KW quantitation; restorative therapy; polymorphic.  
XX  
OS Homo sapiens.  
XX  
PN WO200140521-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 30-NOV-2000; 2000WO-US32758.  
XX  
PR 30-NOV-1999; 99US-0168138.  
PR 29-NOV-2000; 2000US-0726173.  
XX  
PA (CURA-) CURAGEN CORP.  
XX  
PI Shimkets RA, Leach M;  
XX  
DR WPI; 2001-356160/37.

XX immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KW neuroprotective; antimicrobial; gene therapy; vaccine; amyase; cancer;  
XX amyloid protein; angiotensin; apoptosis related protein; cadherin;  
XX cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KW complement related protein; cytochrome; kinesin; cytokine; interferon;  
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KW multifactorial disease; autoimmune disease; infection;  
KW nervous system disease.  
XX  
OS Homo sapiens.  
XX  
PN WO200147944-A2.  
XX  
PD 05-JUL-2001.  
XX  
PF 28-DEC-2000; 2000WO-US35498.  
XX  
PR 28-DEC-1999; 99US-0173419.  
PR 27-DEC-2000; 2000US-0173419.  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX Shinkets RA, Leach M;  
PI  
XX WPI: 2001-465210/50.  
DR  
XX Polymorphic nucleic acids encoding e.g. amyases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g.  
PT cancer, autoimmune diseases and infections -  
XX  
XX Disclosure; Page 4045; 4143pp; English.  
PS  
XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)  
CC encoding polymorphic variants of proteins related to amyases, amyloid  
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,  
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,  
CC complement related proteins, cytochromes, kinesins, cytokines,  
CC interleukins, interlukins, G-protein coupled receptors and thioesterases.  
CC The present sequence is a peptide encoded by one such oligonucleotide.  
CC The oligonucleotides and the peptides encoded by them may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate expression of the proteins listed above. Disorders that may  
CC be prevented, diagnosed and/or treated include multifactorial diseases  
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid  
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus  
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,  
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous  
CC system and an infection of pathogenic organisms.  
CC  
XX  
SQ Sequence 1 AA:  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 X 1  
DB 1 L 1  
RESULT 11  
AAM53218  
ID AAM53218 standard; Peptide: 1 AA.  
XX  
AC AAM53218;  
XX  
DT 09-NOV-2001 (first entry)  
XX  
DE Human nonconservative amino acid changing SNP related peptide SEQ:6913.  
XX  
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
KW protein therapy; vaccine; probe; diagnostic assay; detection;  
XX

KW quantitation; restorative therapy; polymorphic.  
XX  
OS Homo sapiens.  
XX  
PN WO200140521-A2.  
XX  
PD 07-JUN-2001.  
XX  
PF 30-NOV-2000; 2000WO-US32758.  
XX  
PR 30-NOV-1999; 99US-0168138.  
PR 29-NOV-2000; 2000US-0726173.  
XX  
XX (CURA-) CURAGEN CORP.  
PA  
XX Shinkets RA, Leach M;  
PI  
XX WPI: 2001-356160/37.  
DR  
XX Polymorphic nucleic acid sequences, useful in genetic testing and  
PT therapy -  
PT  
XX  
XX Claim 29; Page 2619; 2653pp; English.  
PS  
XX AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide  
CC sequences (I), which contain single nucleotide polymorphisms (SNPs).  
CC AAM53114 to AAM53329 represent peptides related to human polymorphic  
CC polynucleotide sequences. The sequences can be used in gene and protein  
CC therapy, and in vaccine production. (I) and the polypeptides encoded by  
CC them may be used in the prevention, diagnosis and treatment of diseases  
CC associated with inappropriate expression of polymorphic polypeptides.  
CC For example, (I) may be used to treat disorders by rectifying mutations  
CC or deletions in a patient's genome that affect the activity of  
CC polypeptides by expressing inactive proteins or to supplement the  
CC patients own production of polypeptide. Additionally, (I) and its  
CC complementary sequences may also be used as DNA probes in diagnostic  
CC assays to detect and quantitate the presence of similar nucleic acids  
CC in samples, and therefore which patients may be in need of restorative  
CC therapy. The polypeptides encoded by (I) may be used as antigens in the  
CC production of antibodies specific for polymorphic polypeptides. The  
CC antibodies may also be used to down regulate expression and activity.  
CC The antibodies may also be used as diagnostic agents for detecting the  
CC presence of polymorphic polypeptides in samples.  
CC  
XX  
SQ Sequence 1 AA:  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred.No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
OY 1 X 1  
DB 1 H 1  
RESULT 12  
AAM53219  
ID AAM53219 standard; Peptide: 1 AA.  
XX  
AC AAM53219;  
XX  
DT 09-NOV-2001 (first entry)  
XX  
DE Human nonconservative amino acid changing SNP related peptide SEQ:6914.  
XX  
KW Human; single nucleotide polymorphism; SNP; genome; gene therapy;  
KW protein therapy; vaccine; probe; diagnostic assay; detection;  
KW quantitation; restorative therapy; polymorphic.  
XX  
OS Homo sapiens.  
XX  
PN WO200140521-A2.  
XX

XX	AAM97974;
AC	
DT	24-JAN-2002 (first entry)
XX	
DE	Human peptide #1249 encoded by a SNP oligonucleotide.
XX	
KW	Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KM	neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW	amyloid protein; angiotensinogen; apoptosis related protein; cadherin;
KM	cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW	complement related protein; cytochrome; kinesin; cytokine; interferon;
KM	interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW	multifactorial disease; autoimmune disease; infection;
KM	nervous system disease.
OS	Homo sapiens.
XX	
PN	WO200147944-A2.
XX	
PD	05-JUL-2001.
XX	
PF	28-DEC-2000; 2000MO-US35498.
XX	
PR	28-DEC-1999; 99US-0173419.
PR	27-DEC-2000; 2000US-0173419.
XX	(CURA-) CURAGEN CORP.
PA	
XX	
PI	Shinkets RA, Leach M;
XX	
DR	WPI: 2001-465210/50.
PT	polymorphic nucleic acids encoding e.g. amyloses, cyclins, polymerases,
PR	oncogenes and histones, useful for diagnosing and treating, e.g.
PS	cancer, autoimmune diseases and infections -
XX	
PS	Disclosure: Page 3941; 4143pp; English.
XX	
CC	The present invention relates to oligonucleotides (see AAL26793-AAI34659)
CC	encoding polymorphic variants of proteins related to amylases, amyloid
CC	proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,
CC	polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC	complement related proteins, cytochromes, kinesins, cytokines,
CC	interferons, interleukins, G-protein coupled receptors and thioesterases.
CC	The present sequence is a peptide encoded by one such oligonucleotide.
CC	The oligonucleotides and the peptides encoded by them may be used in the
CC	detection, diagnosis and treatment of disorders associated with
CC	inappropriate expression of the proteins listed above. Disorders that may
CC	be prevented, diagnosed and/or treated include multifactorial diseases
CC	with a genetic component, such as autoimmune diseases (e.g. Rheumatoid
CC	arthritis), multiple sclerosis, diabetes, systemic lupus erythematosus
CC	and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC	brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC	system and an infection of pathogenic organisms.
SQ	Sequence 1 AA:
OY	Query Match 0.0%; Score 0; DB 22; Length 1; Best Local Similarity 0.0%; Pred. No. 0; Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
DB	1 L 1
RESULT 9 ID AAM98354 standard; Peptide: 1 AA. AC AAM98354; XX	

DT		24-JAN-2002	(first entry)
XX			
DE		Human peptide #1629 encoded by a SNP oligonucleotide.	
XX			
KW		Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;	
KM		neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;	
KV		amyloid protein; angiotensin; apoptosis related protein; cadherin;	
KW		cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;	
KV		complement related protein; cytochrome; kinesin; cytokine; interferon;	
KM		interleukin; G-protein coupled receptor; thioesterase; inflammation;	
KW		multifactorial disease; autoimmune disease; infection;	
KV		nervous system disease.	
XX			
OS		Homo sapiens.	
PN		WO200147944-A2.	
XX			
PD		05-JUL-2001.	
XX			
PF		28-DEC-2000; 2000WO-US35498.	
XX			
PR		28-DEC-1999; 99US-0173419.	
PR		27-DEC-2000; 2000US-0173419.	
XX			
PA		(CURA-) CURAGEN CORP.	
XX			
PI		Shinkets RA, Leach M;	
XX			
DR		WPI; 2001-465210/50.	
XX			
PT		Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,	
PT		oncogenes and histones, useful for diagnosing and treating, e.g.	
PT		cancer, autoimmune diseases and infections -	
PS		Disclosure; Page 4025; 4143p; English.	
CC		The present invention relates to oligonucleotides (see AML26793-AU34659)	
CC		encoding polymorphic variants of proteins related to amylases, amyloid	
CC		proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,	
CC		polymerase, oncogenes, histones, kinases, colony stimulating factors,	
CC		complement related proteins, cytochromes, kinesins, cytokines,	
CC		interferons, interleukins, G-protein coupled receptors and thioesterases.	
CC		The present sequence is a peptide encoded by one such oligonucleotide.	
CC		The oligonucleotides and the peptides encoded by them may be used in the	
CC		prevention, diagnosis and treatment of diseases associated with	
CC		inappropriate expression of the proteins listed above. Disorders that may	
CC		be prevented, diagnosed and/or treated include multifactorial diseases	
CC		with a genetic component, such as autoimmune diseases (e.g. rheumatoid	
CC		arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus	
CC		and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,	
CC		breast, colon and kidney, leukaemia), diseases of the nervous	
CC		system and an infection of pathogenic organisms.	
XX			
SO		Sequence 1 AA:	
OY	1 X 1		
DB	1 S 1		
		Query Match 0.0%, Score 0; DB 22; Length 1;	
		Best Local Similarity 0.0%; Pwd. No. 0;	
		Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
		RESULT 10	
ID	AAAG8447		
	AAAG8447 standard; Peptide: 1 AA.		
XX			
AC	AAAG8447;		
XX			
DT	24-JAN-2002	(first entry)	
XX			
DE		Human peptide #1722 encoded by a SNP oligonucleotide.	

Db 1 M 1

RESULT 6  
AAM97643

ID AAM97643 standard; Peptide: 1 AA.

XX AAM97643;

DT 24-JAN-2002 (first entry)

XX Human peptide #918 encoded by a SNP oligonucleotide.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KM neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
KM amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KM cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KM complement related protein; cytochrome; kinesin; cytokine; interferon;  
KM interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KM multifactorial disease; autoimmune disease; infection;  
KM nervous system disease.

OS Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

PA Shinkets RA, Leach M;

DR WPI: 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g.  
PT cancer, autoimmune diseases and infections -

PS Disclosure: Page 3869; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)  
CC encoding polymorphic variants of proteins related to amylases, amyloid  
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,  
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,  
CC complement related proteins, cytochromes, kinesins, cytokines,  
CC interleukins, G-protein coupled receptors and thioesterases.  
CC The present sequence is a peptide encoded by one such oligonucleotide.  
CC The oligonucleotides and the peptides encoded by them may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate expression of the proteins listed above. Disorders that may  
CC be prevented, diagnosed and/or treated include multifactorial diseases  
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid  
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus  
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,  
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous  
CC system and an infection of pathogenic organisms.

SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1

Db 1 C 1

RESULT 7  
AAM97834  
ID AAM97834 standard; Peptide: 1 AA.

XX AAM97834;

DT 24-JAN-2002 (first entry)

XX Human peptide #1109 encoded by a SNP oligonucleotide.

XX Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;  
KM neuroprotective; antimicrobial; gene therapy; vaccine; amylose; cancer;  
KM amyloid protein; angiotensin; apoptosis related protein; cadherin;  
KM cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;  
KM complement related protein; cytochrome; kinesin; cytokine; interferon;  
KM interleukin; G-protein coupled receptor; thioesterase; inflammation;  
KM multifactorial disease; autoimmune disease; infection;  
KM nervous system disease.

OS Homo sapiens.

PN WO200147944-A2.

PD 05-JUL-2001.

PF 28-DEC-2000; 2000WO-US35498.

PR 28-DEC-1999; 99US-0173419.

PR 27-DEC-2000; 2000US-0173419.

XX (CURA-) CURAGEN CORP.

PA Shinkets RA, Leach M;

DR WPI: 2001-465210/50.

XX Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,  
PT oncogenes and histones, useful for diagnosing and treating, e.g.  
PT cancer, autoimmune diseases and infections -

PS Disclosure: Page 3911; 4143pp; English.

XX The present invention relates to oligonucleotides (see AAL26793-AAL34659)  
CC encoding polymorphic variants of proteins related to amylases, amyloid  
CC proteins, angiotensin, apoptosis related proteins, cadherin, cyclin,  
CC polymerase, oncogenes, histones, kinases, colony stimulating factors,  
CC complement related proteins, cytochromes, kinesins, cytokines,  
CC interleukins, G-protein coupled receptors and thioesterases.  
CC The present sequence is a peptide encoded by one such oligonucleotide.  
CC The oligonucleotides and the peptides encoded by them may be used in the  
CC prevention, diagnosis and treatment of diseases associated with  
CC inappropriate expression of the proteins listed above. Disorders that may  
CC be prevented, diagnosed and/or treated include multifactorial diseases  
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid  
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythematosus  
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,  
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous  
CC system and an infection of pathogenic organisms.

SQ Sequence 1 AA;

Query Match 0.0%; Score 0; DB 22; Length 1;

Best Local Similarity 0.0%; Pred. No. 0;

Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1

Db 1 E 1

RESULT 8  
AAM97974  
ID AAM97974 standard; Peptide: 1 AA.

CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB5737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1 AA:  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1  
Db 1 H 1  
RESULT 4  
ABB66810  
ID ABB66810 standard; Protein; 1 AA.  
XX  
AC ABB66810;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 27222.  
XX  
KW Drosophila; developmental biology; cell signalling; insecticide;  
KM pharmaceutical.  
XX  
OS Drosophila melanogaster.  
XX  
PN WO200171042-A2.  
XX  
PD 27-SEP-2001.  
XX  
PF 23-MAR-2001; 2001WO-US09231.  
XX  
PR 23-MAR-2000; 2000US-191637P.  
XX  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PI Venter JC, Adams M, Li PWD, Myers EW;  
XX  
DR WPI: 2001-656860/75.  
XX  
DR N-PSDB; ABL10913.  
XX  
PT New isolated nucleic acid detection reagent for detecting 1000 or more  
PT genes from Drosophila and for elucidating cell signalling and cell-cell  
PT interactions -  
XX  
PS Disclosure; SEQ ID NO 27222; 21pp + Sequence Listing; English.  
XX  
XX The invention relates to an isolated nucleic acid detection reagent  
CC capable of detecting 1000 or more genes from Drosophila. The invention is  
CC useful in developmental biology and in elucidating cell signalling and  
CC cell-cell interactions in higher eukaryotes for the development of  
CC insecticides, therapeutics and pharmaceutical drugs. The invention  
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA  
CC sequences (ABL01840-ABL16175) and the encoded proteins  
CC (ABB5737-ABB72072).  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1 AA:  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 X 1  
Db 1 H 1  
RESULT 5  
ABG02941  
ID ABG02941 standard; Protein; 1 AA.  
XX  
AC ABG02941;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #2932.  
XX  
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
XX  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI: 2001-639362/73.  
XX  
DR N-PSDB; AAS67128.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity -  
XX  
PS Claim 20; SEQ ID NO 33300; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping, and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 1 AA:  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 X 1



PT New immunogenic peptides with HLA binding motif, useful in treatment  
PT and diagnosis of cancers and viral diseases -  
PS Claim 1; Page 80; 150pp; English.  
XX  
CC AA145390 to AA148214 represent specifically claimed immunogenic peptides  
CC having a human major histocompatibility complex (MHC) Class I (also  
CC known as human leukocyte antigen (HLA)) binding motif. The immunogenic  
CC peptides can bind to a specific HLA allele (i.e. HLA-A subtypes  
CC HLA-A2.1, A1, A3.2 or A24.1 or HLA-B or C) and induce a cytotoxic T cell  
CC response against the antigen from which the peptide is derived.  
CC Cytotoxic T lymphocytes (CTLs) which destroy antigen-bearing cells are  
CC normally induced by an antigen in the form of a peptide fragment bound  
CC to a HLA molecule, rather than the intact foreign antigen itself, and  
CC are particularly important in tumour rejection and in fighting viral  
CC infections. The peptides are therefore useful therapeutically to treat  
CC or prevent viral infections and cancers in mammals (especially humans)  
CC e.g. prostate cancer, hepatitis B and C, AIDS, and renal carcinoma.  
CC They can be administered as vaccines to elicit an immune response in  
CC individuals susceptible or otherwise at risk of viral infection or  
CC cancer, or used to treat chronic or acute conditions. They are also  
CC useful diagnostically, and can be used to induce a cytotoxic T cell  
CC response, by contacting a cytotoxic T cell with the peptide e.g. to  
CC produce CTLs ex vivo for infusion back into a patient. The  
CC polynucleotides encoding the immunogenic peptides are also useful  
CC therapeutically and for immunisation as above.  
SQ Sequence 1 AA;  
XX  
Query Match 0.0%; Score 0; DB 20; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 x 1  
Db 1 x 1  
XX  
RESULT 2  
ABBS6870  
ID ABBS6870 standard; Peptide; 1 AA.  
XX  
AC ABBS6870;  
XX  
DT 05-MAR-2002 (first entry)  
XX  
DE Human SNP related amino acid sequence SEQ ID NO:1435.  
XX  
KW Human; single nucleotide polymorphism; SNP; polymorphism; cytostatic;  
KW immunosuppressive; antiinflammatory; neuroprotective; antimicrobial;  
KW autoimmune disease; inflammation; cancer; nervous system disease;  
KW infection; polymorphic protein.  
XX  
OS Homo sapiens.  
XX  
OS WO200138586-A2.  
XX  
PN 31-MAY-2001.  
XX  
PD 22-NOV-2000; 2000WO-US32311.  
XX  
PF 24-NOV-1999; 99US-0167383.  
XX  
PR (CURA-) CURAGEN CORP.  
XX  
PA Shimkets RA, Leach M;  
XX  
PI WPI: 2001-355949/37.  
XX  
DR Isolated human nucleic acids comprising one or more single nucleotide  
XX polymorphisms, useful for treating a subject suffering from a  
XX pathology, e.g. autoimmune diseases, ascribed to the presence of a  
XX sequence polymorphism -

XX  
PS Claim 1; Page 666; 674pp; English.  
XX  
CC ABL00010 to ABL01104 represent human nucleic acid oligonucleotides  
CC comprising one or more single nucleotide polymorphisms (SNPs). ABB56531  
CC to ABB56903 represent human peptides encoded by some of the SNP  
CC oligonucleotides. The sequences from the present invention can have  
CC immunosuppressive, cytostatic, antiinflammatory, neuroprotective and  
CC antimicrobial activities. Nucleic acids, polypeptides, oligonucleotides  
CC and antibodies from the present invention can be used for treating a  
CC subject suffering from, at risk for, or suspected of, suffering from a  
CC pathology ascribed to the presence of a sequence polymorphism. The  
CC pathology may be autoimmune diseases, inflammation, cancer, diseases of  
CC the nervous system, and infection by pathogenic microorganisms. The SNPs  
CC are also useful for determining which forms of a characterised  
CC polymorphism are present in individuals. The antibodies may be used in  
CC the detection, quantitation and/or cellular or tissue localisation of a  
CC polymorphic protein (e.g., for use in measuring levels of the  
CC polymorphic protein within appropriate physiological samples).  
SQ Sequence 1 AA;  
XX  
Query Match 0.0%; Score 0; DB 22; Length 1;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 x 1  
Db 1 s 1  
XX  
RESULT 3  
ABB66809  
ID ABB66809 standard; Protein; 1 AA.  
XX  
AC ABB66809;  
XX  
DT 26-MAR-2002 (first entry)  
XX  
DE Drosophila melanogaster polypeptide SEQ ID NO 27219.  
XX  
DE Drosophila; developmental biology; cell signalling; insecticide;  
KW pharmaceutical.  
XX  
KW Drosophila melanogaster.  
XX  
OS WO200171042-A2.  
XX  
PN 27-SEP-2001.  
XX  
PD 23-MAR-2001; 2001WO-US09231.  
XX  
PF 23-MAR-2000; 2000US-191637P.  
XX  
PR 11-JUL-2000; 2000US-0614150.  
XX  
PA (PEKE ) PE CORP NY.  
XX  
PN Venter JC, Adams M, Li PWD, Myers EW;  
XX  
PI WPI: 2001-656860/75.  
XX  
DR N-PSDB: ABL10912.  
XX  
DR New isolated nucleic acid detection reagent for detecting 1000 or more  
XX genes from Drosophila and for elucidating cell signalling and cell-cell  
XX interactions -  
XX  
PS Disclosure; SEQ ID NO 27219; 21pp + Sequence Listing; English.  
XX  
CC The invention relates to an isolated nucleic acid detection reagent  
XX capable of detecting 1000 or more genes from Drosophila. The invention is  
XX useful in developmental biology and in elucidating cell signalling and  
XX cell-cell interactions in higher eukaryotes for the development of  
XX insecticides, therapeutics and pharmaceutical drugs. The invention

GenCore version 5.1.4\_p5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:27:06 ; Search time 5 Seconds  
(without alignments)  
186.551 Million cell updates/sec

Title: US-09-877-606-8  
Perfect score: 7  
Sequence: 1 xxxxxxx 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues  
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_101002.\*  
1: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT.\*  
8: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT.\*  
9: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT.\*  
10: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT.\*  
11: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT.\*  
12: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT.\*  
13: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT.\*  
14: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT.\*  
15: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT.\*  
16: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT.\*  
17: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT.\*  
18: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT.\*  
19: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT.\*  
20: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT.\*  
21: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT.\*  
22: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*  
23: /SIDSL/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	0	0.0	1	20	AAV46652
2	0	0.0	1	22	ABB56870
3	0	0.0	1	22	ABB66809
4	0	0.0	1	22	ABB66810
5	0	0.0	1	22	ABC02941
6	0	0.0	1	22	AAW97643
7	0	0.0	1	22	AAW97834
8	0	0.0	1	22	AAW97974
9	0	0.0	1	22	AAW98354
10	0	0.0	1	22	AAW98447

11	0	0.0	1	22	AAM53218	Human nonconservat
12	0	0.0	1	22	AAM53219	Human nonconservat
13	0	0.0	1	22	AAM53290	Human nonconservat
14	0	0.0	1	22	AAM53291	Human nonconservat
15	0	0.0	1	22	AAM53328	Human nonconservat
16	0	0.0	1	22	AAM53329	Human nonconservat
17	0	0.0	1	22	AAM45230	H11 binding site c
18	0	0.0	1	22	AAG99966	ERA binding domain
19	0	0.0	1	22	AAG99983	ERA binding domain
20	0	0.0	1	22	AAG99987	ERA binding domain
21	0	0.0	1	22	AAG99988	ERA binding domain
22	0	0.0	1	22	AAM00010	ERA binding domain
23	0	0.0	1	22	AAM00011	ERA binding domain
24	0	0.0	1	22	AAM00013	ERA binding domain
25	0	0.0	1	22	AAM00016	ERA binding domain
26	0	0.0	1	22	AAG98026	Human SNP associat
27	0	0.0	1	22	AAG98134	Human SNP associat
28	0	0.0	1	22	AAB91029	Thyrotropin releas
29	0	0.0	1	22	AAB91546	Endothelins and re
30	0	0.0	1	22	AAB91665	Opioid peptide SEQ
31	0	0.0	1	22	AAB91739	Opioid peptide SEQ
32	0	0.0	1	22	AAB91892	Apoptosis related
33	0	0.0	1	22	AAB92150	Polypeptide SEQ ID
34	0	0.0	1	22	AAB92392	Miscellaneous pept
35	0	0.0	1	23	ABG63437	Human albumin fusi
36	0	0.0	1	23	ABG63439	Human albumin fusi
37	0	0.0	1	23	ABG63578	Human albumin fusi
38	0	0.0	1	23	ABG63753	Human albumin fusi
39	0	0.0	1	23	ABG63997	Human albumin fusi
40	0	0.0	1	23	ABG64298	Human albumin fusi
41	0	0.0	1	23	ABG64370	Human albumin fusi
42	0	0.0	1	23	ABG64797	Human albumin fusi
43	0	0.0	1	23	ABB78513	GAGP hydroxyprolin
44	0	0.0	1	23	ABB78514	GAGP hydroxyprolin
45	0	0.0	1	23	ABB78515	GAGP hydroxyprolin

ALIGNMENTS

RESULT 1  
ID AAY46652 standard; Peptide: 1 AA.  
AC AAY46652;  
XX  
DT 01-DEC-1999 (first entry)  
XX  
XX Immunogenic peptide having a human leukocyte antigen binding motif #1263.  
DE  
XX Human leukocyte antigen; binding; immunogenic; glycoprotein; MHC; HLA;  
KW Immune response; T cell activation; major histocompatibility complex;  
KW cytotoxic T lymphocyte; CTL; tumour rejection; viral infection; cancer;  
KW prostate cancer; hepatitis B; hepatitis C; AIDS; renal carcinoma;  
KW vaccine; immunisation.  
XX  
OS Synthetic.  
XX Homo sapiens.  
XX  
PN WO9945954-A1.  
PD 16-SEP-1999.  
XX  
PF 13-MAR-1998; 98WO-US05039.  
XX  
PR 13-MAR-1998; 98WO-US05039.  
XX  
PA (EPIM-) EPIMUNE INC.  
XX  
PI Sette A, Kubo RT, Sidney J, Celis E, Grey HM, Southwood S;  
XX WPI: 1999-551214/46.  
XX

TITLE  
JOURNAL  
COMMENT  
Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estevatson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccleresgen.com  
Insert Length: 1185 Std Error: 0.00  
Seq primer: -40RP from Gibco  
High quality sequence stop: 396.  
Location/Qualifiers

## FEATURES

source

1. 437  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1015-1805"  
/clone\_lib="Gm-c1015"  
/tissue\_type="Mature flowers, field grown plants"  
/lab\_host="XL10-Gold"  
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from mature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR cDNA library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 144 a 69 c 119 g 105 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.22e+03 Length: 437  
Score: 44.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 91.67% Indels: 0  
DB: 10 Gaps: 0

US-09-877-606-7 (1-7) x AW156192 (1-437)

Oy 1 LeutPAsnTRPpheasn 6  
Db 398 CTCTGGAATGCTTAAAC 381

RESULT 15  
AW458604 437 bp mRNA linear EST 03-DEC-2001  
LOCUS sh1c05.y1 Gm-c1016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
DEFINITION Gm-c1016-4449 5', mRNA sequence.  
ACCESSION AW458604  
VERSION AW458604.1 GI:7028821  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

REFERENCE 1 (bases 1 to 437)  
AUTHORS Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V., Khanna

TITLE  
JOURNAL  
COMMENT  
A., Bolla, B., Maria, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: estevatson.wustl.edu

This clone is available through: ResGen, Invitrogen Corp. 2130 South Memorial Parkway Huntville, AL 35801 For further information call: (800)-533-4363 or contact via email: ccleresgen.com  
Insert Length: 948 Std Error: 0.00  
Seq primer: -40RP from Gibco  
High quality sequence stop: 387.  
Location/Qualifiers

## FEATURES

source

1. 437  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-4449"  
/clone\_lib="Gm-c1016"  
/tissue\_type="Immature flowers of field grown plants"  
/lab\_host="XL10-Gold"  
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2: XhoI; This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 163 a 91 c 50 g 133 t  
ORIGIN

Alignment Scores:  
Pred. No.: 3.22e+03 Length: 437  
Score: 44.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 91.67% Indels: 0  
DB: 10 Gaps: 0

US-09-877-606-7 (1-7) x AW458604 (1-437)

Oy 1 LeutPAsnTRPpheasn 6  
Db 116 CTCTGGAATGCTTAAAC 99

Search completed: March 25, 2003, 17:27:55  
Job time : 200.31 secs

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
Insert Length: 1143 Std Error: 0.00  
Seq primer: -408P from Gibco  
High quality sequence stop: 428.

FEATURES  
source

1.429  
Location/Qualifiers  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1017-4758"  
/clone\_1lb="Gm-c1017"  
/tissue\_type="vegetable buds of field grown plants"  
/lab\_host="X10-Gold"  
/note="Vector: Bluescript II XR, Site\_1: EcoRI, Site\_2: XhoI; This cDNA library was constructed from mRNA isolated from vegetable buds of field grown plants. The cDNA library was prepared using the Stratagene Bluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into X10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 130 a 91 c 77 g 131 t  
ORIGIN

Alignment Scores:

Pred. No.: 3.14e+03 Length: 429  
Score: 44.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 91.67% Indels: 0  
DB: 10 Gaps: 0

US-09-877-606-7 (1-7) x AM395321 (1-429)

QY 1 LeuTrpAsnTrpPheAsn 6

Db 298 CTTTGGAATTGGTTTAC 281

RESULT 13

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

soybean.

Glycine max

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 436)

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

,Y., Person,B., Swaller,T., Gibbons,M., Page,D., Harvey,N., Schurk

,R., Rlter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann

TITLE  
JOURNAL  
COMMENT

,R., Waterston,R. and Wilson,R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: ccu@resgen.com  
Insert Length: 1105 Std Error: 0.00  
High quality sequence stop: 407.

FEATURES  
source

1.436  
Location/Qualifiers  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1068-1223"  
/clone\_1lb="Gm-c1068"  
/tissue\_type="leaf, drought stressed, 1 month old plants, greenhouse grown"  
/lab\_host="DH10B"  
/note="Vector: Bluescript II SK+, Site\_1: EcoRI, Site\_2: XhoI; The cDNA library was constructed from mRNA isolated from drought stressed leaf tissue of the cultivar Williams 82. The month old greenhouse grown plants were deprived of water for 3 days prior to harvesting the stressed leaf tissue. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

BASE COUNT 133 a 90 c 77 g 136 t  
ORIGIN

Alignment Scores:

Pred. No.: 3.21e+03 Length: 436  
Score: 44.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 91.67% Indels: 0  
DB: 12 Gaps: 0

US-09-877-606-7 (1-7) x BF324158 (1-436)

QY 1 LeuTrpAsnTrpPheAsn 6

Db 296 CTTTGGAATTGGTTTAC 279

RESULT 14

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

soybean.

Glycine max

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eustosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;

Glycine.

1 (bases 1 to 437)

Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna

,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,

Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers

## AUTHORS

Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y., and Shinozaki, K.

## TITLE

Large scale analysis of Arabidopsis full-length cDNA (2002b)

## COMMENT

Unpublished (2002)  
Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan  
Tel: 81-298-36-4359  
Fax: 81-298-36-9060

Email: mseki@riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambda PhiC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified Bluescript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

## FEATURES

## source

Location/Qualifiers  
1..409  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/clone="RAFL09-20-H08"  
/clone\_11b="RAFL9"  
/dev\_stage="plants at various developmental stages from germination to mature seeds"  
/lab\_host="DH10B"  
/note="Site\_1: BamHI; Site\_2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments 64 g 121 t

## BASE COUNT

139 a 85 c 64 g 121 t

## ORIGIN

## Alignment Scores:

Pred. No.: 2.94e+03 Length: 409  
Score: 44.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 91.67% Indels: 0  
DB: 10 Gaps: 0

US-09-877-606-7 (1-7) x AV799546 (1-409)

Oy 1 LeuTPAsnTrpPheAsn 6

Db 277 TTATGCACTGCTCAAT 294

## RESULT 11

## LOCUS

AM306467/c 418 bp mRNA linear EST 02-DEC-2001  
se51b07.y1 Gm-c1017 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-c1017-2414 5', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

AM306467.1 GI:6718820  
EST.  
soybean.  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

## REFERENCE

## AUTHORS

1 (bases 1 to 418)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project  
Unpublished (1999)  
Contact: Shoemaker R/Public Soybean EST Project

## TITLE

JOURNAL COMMENT

Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810

Email: est@watson.wustl.edu  
This clone is available through: ResGen, Invitrogen Corp. 2130  
South Memorial Parkway Huntsville, AL 35801 For further information  
call: (800)-533-4363 or contact via email: [custresgen.com](mailto:custresgen.com)  
Insert length: 1042 Std Error: 0.00  
Seq primer: -40RP from Gibco  
High quality sequence stop: 375.

## FEATURES

## source

Location/Qualifiers  
1..418  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-c1017-2414"  
/clone\_11b="Gm-c1017"  
/tissue\_type="vegetable buds of field grown plants"  
/lab\_host="XL10-Gold"  
/note="Vector: pBluescript II XR; Site\_1: EcoRI; Site\_2: XhoI; This cDNA library was constructed from mRNA isolated from vegetable buds of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

## BASE COUNT

125 a 90 c 75 g 128 t

## ORIGIN

## Alignment Scores:

Pred. No.: 3.03e+03 Length: 418  
Score: 44.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 91.67% Indels: 0  
DB: 10 Gaps: 0

US-09-877-606-7 (1-7) x AM306467 (1-418)

Oy 1 LeuTPAsnTrpPheAsn 6

Db 298 CTTTGGAAATGGTTTAAAC 281

## RESULT 12

## LOCUS

AM395321 429 bp mRNA linear EST 03-DEC-2001  
SH46f03.y1 Gm-c1017 glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
Gm-c1017-4758 5', mRNA sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

AM395321.1 GI:5913791  
EST.  
soybean.  
Glycine max  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
Glycine.

## REFERENCE

## AUTHORS

1 (bases 1 to 429)  
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Stepien, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.  
Public Soybean EST Project

## TITLE

RESULT 8  
 A1938434/c 327 bp mRNA linear EST 30-NOV-2001  
 LOCUS sc45b12.y1 Gm-cl015 GLYCINE max cDNA clone GENOME SYSTEMS CLONE ID:  
 DEFINITION Gm-cl015-504 5', mRNA sequence.  
 ACCESSION A1938434  
 VERSION A1938434.1 GI:5677304  
 KEYWORDS EST.  
 SOURCE soybean.  
 ORGANISM Glycine max  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;  
 Glycine.  
 REFERENCE 1 (bases 1 to 327)  
 AUTHORS Shoemaker, R., Kelm, P., Vodka, L., Erpelting, J., Corryell, V., Khanna  
 'A', Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, D., Beck, C.,  
 Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers  
 'Y', Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk  
 'R', Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann  
 'R', Mattern, R. and Wilson, R.  
 TITLE Public Soybean EST Project  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Shoemaker R/Public Soybean EST Project  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available through: Resgen, Invitrogen Corp. 2130  
 South Memorial Parkway Huntville, AL 35801 For further information  
 call: (800)-533-4363 or contact via email: cch@resgen.com  
 Seq primer: -40RP from G1bco  
 High quality sequence stop: 321.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..327  
 /organism="Glycine max"  
 /db\_xref="taxon:3847"  
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl015-504"  
 /clone\_lib="Gm-cl015"  
 /tissue\_type="Mature flowers, field grown plants"  
 /lab\_host="X110-Gold"  
 /note="Vector: pBluescript II Xr; Site:1: EcoRI; Site:2:  
 XhoI; This cDNA library was constructed from mRNA isolated  
 from mature flowers of field grown plants. The cDNA  
 library was prepared using the Stratagene pBluescript II  
 XR cDNA library construction kit. Complementary DNA was  
 synthesized from mRNA using a primer consisting of a poly  
 (dT) sequence with a XhoI restriction site. EcoRI adapters  
 were ligated to the blunt-ended cDNA fragments followed by  
 XhoI digestion. The cDNA fragments were directionally  
 cloned into the EcoRI-XhoI restriction site of the  
 pBluescript vector. The ligated cDNA fragments were  
 transformed into X110-Gold host cells. This library was  
 constructed by Dr. Randy Shoemaker and Dr. John  
 Erpelting."  
 BASE COUNT 67 a 87 c 71 g 102 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.15e+03 Length: 327  
 Score: 44.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 91.67% Indels: 0  
 Gaps: 0  
 DB: 9  
 US-09-877-606-7 (1-7) x A1938434 (1-327)  
 Oy 2 TrpAsnTrpPheAsnIle 7  
 ||||||||||||||||

Db 297 TCGAAGCTGCTTCACATC 280  
 RESULT 9  
 B6604081 376 bp mRNA linear EST 14-AUG-2001  
 LOCUS EST503171 Plasmodium yoelii sporozoite cDNA Plasmodium yoelii cDNA  
 DEFINITION clone pYCDX81, mRNA sequence.  
 ACCESSION B6604081  
 VERSION B6604081.1 GI:15154095  
 KEYWORDS EST.  
 SOURCE Plasmodium yoelii.  
 ORGANISM Plasmodium yoelii  
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
 REFERENCE 1 (bases 1 to 376)  
 AUTHORS Kappe, S.H.I., Gardner, M.J., Brown, S.M., Ross, J., Matuschewski, K.,  
 Ribeiro, J.M., Adams, J.H., Quekenbush, J., Cho, J., Carucci, D.J.,  
 Hoffman, S.L. and Nussenzweig, V.  
 TITLE Exploring the transcriptome of the malaria sporozoite stage  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (17), 9895-9900 (2001)  
 COMMENT Contact: Malcolm J. Gardner  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: gardner@tigr.org  
 Request for clones, please contact: Stefan Kappe,  
 kappe@leopold.mll.med.nyu.edu Michael Heidelberger Division,  
 Department of Pathology New York University School of Medicine.  
 FEATURES  
 source  
 Location/Qualifiers  
 1..376  
 /organism="Plasmodium yoelii"  
 /strain="17XNL"  
 /db\_xref="taxon:5861"  
 /clone\_lib="pYCDX81"  
 /clone="Plasmodium yoelii sporozoite cDNA"  
 /dev\_stage="sporozoites from salivary gland"  
 /lab\_host="E. coli TOP10"  
 /note="Vector: pCR4; TA cloning; Plasmodium yoelii  
 sporozoite cDNA library from salivary gland sporozoites 14  
 days post-infection"  
 BASE COUNT 158 a 40 c 52 g 126 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.61e+03 Length: 376  
 Score: 44.00 Matches: 6  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 91.67% Indels: 0  
 Gaps: 0  
 DB: 12  
 US-09-877-606-7 (1-7) x B6604081 (1-376)  
 Oy 2 TrpAsnTrpPheAsnIle 7  
 ||||||||||||||||  
 Db 165 TCGATTGCTTAATATT 182  
 RESULT 10  
 AV799546 409 bp mRNA linear EST 29-MAR-2002  
 LOCUS AV799546 RAF19 Arabidopsis thaliana cDNA clone RAF109-20-H08 3',  
 DEFINITION mRNA sequence.  
 ACCESSION AV799546  
 VERSION AV799546.1 GI:19833531  
 KEYWORDS EST.  
 SOURCE thale cress.  
 ORGANISM Arabidopsis thaliana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 REFERENCE 1 (bases 1 to 409)

/db\_xref="taxon:3702"  
/clone="T7N18"  
/clone\_lib="TAMU"  
/sex="hermaphrodite"  
/note="Vector: BelOBACII; Site\_1: HindIII; Site\_2: HindIII  
; Produced by Rod Wing"

BASE COUNT 145 a 75 c 106 g 218 t

ORIGIN

Alignment Scores:

Pred. No.:	3.12e+03	Length:	544
Score:	45.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	85.71%	Mismatches:	0
Query Match:	93.75%	Indels:	0
DB:	17	Gaps:	0

US-09-877-606-7 (1-7) x B28259 (1-544)

OY 1 LeutpAsnTrpPheAsnIle 7  
|||||

Db 120 CTTGCACTGCTATATATT 140

RESULT 6 647 bp DNA linear GSS 09-DEC-1997

B67686 T23011TF TAMU Arabidopsis thaliana genomic clone T23011, DNA  
LOCUS sequence.

ACCESSION B67686  
VERSION B67686.1 GI:2666440  
KEYWORDS GSS.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
1 (bases 1 to 647)  
Rounsley,S.D., Field,C.E., Bass,S., Linher,K., Linher,K., Golden,K.,  
Berry,K., Granger,D., Suh,E., Wible,C., Adams,M.D. and Venter  
J.C.  
A BAC End Sequence Database for Identifying Minimal Overlaps in  
Arabidopsis Genomic Sequencing. Update 3  
JOURNAL Unpublished (1997)  
COMMENT Other\_GSSs: T23011TR  
Contact: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: rounsley@tigr.org  
Seq primer: M13-21  
Class: BAC ends  
High quality sequence stop: 647.  
Location/Qualifiers  
1. .647  
/organism="Arabidopsis thaliana"  
/strain="Columbia"  
/db\_xref="taxon:3702"  
/clone="T23011"  
/clone\_lib="TAMU"  
/sex="hermaphrodite"  
/note="Vector: BelOBACII; Site\_1: HindIII; Site\_2: HindIII  
; Produced by Rod Wing"

BASE COUNT 193 a 87 c 126 g 240 t 1 others

ORIGIN

Alignment Scores:

Pred. No.:	3.98e+03	Length:	647
Score:	45.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	1
Best Local Similarity:	85.71%	Mismatches:	0
Query Match:	93.75%	Indels:	0

DB: 17 Gaps: 0

US-09-877-606-7 (1-7) x B67686 (1-647)

OY 1 LeutpAsnTrpPheAsnIle 7  
|||||

Db 113 CTTGCACTGCTATATATT 133

RESULT 7 289 bp mRNA linear EST 08-APR-2002

BQ095463 LOCUS  
DEFINITION KK05e10.y1 Ascaris suum female ovary Ascaris suum cDNA 5', mRNA  
sequence.

ACCESSION BQ095463  
VERSION BQ095463.1 GI:20076680  
KEYWORDS EST.  
SOURCE pig roundworm.  
ORGANISM Ascaris suum  
Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea  
; Ascarididae; Ascaris.  
1 (bases 1 to 289)  
McCartter,J., Clifton,S., Chiappelli,B., Page,D., Martin,J., Wylie,T.,  
Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B., Bowers,T.,  
Gibbons,M., Ritter,E., Bennett,J., Franklin,C., Tsagarelisvili,R.,  
Ronko,I., Kennedy,S., Maguire,L., Beck,C., Underwood,K., Steptoe  
,M., Allen,M., Person,B., Swaller,T., Harvey,N., Schurk,R., Kohn,S.,  
Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and  
Wilson,R.  
The Washington Univ. Nematode EST Project, 1999  
Unpublished (1999)  
Contact: McCartter JP  
The Washington Univ. Nematode EST Project, 1999  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
The library was donated by Dr. T. Geary, Pharmacia Inc, Kalamazoo,  
MI, (tgeary@am.pnu.com). DNA Sequencing by: Washington University  
Genome Sequencing Center  
Seq primer: T7 from GIBCO  
High quality sequence stop: 280.  
Location/Qualifiers  
1. .289  
/organism="Ascaris suum"  
/db\_xref="taxon:6253"  
/clone\_lib="Ascaris suum female ovary"  
/sex="Female"  
/tissue\_type="Ovary"  
/dev\_stage="Adult"  
/lab\_host="DH10B"  
/note="Vector: Lambda Zap II(Stratagene); Site\_1: EcoRI;  
Site\_2: EcoRI. The library was donated by Dr. T. Geary,  
Pharmacia Inc, Kalamazoo, MI, (tgeary@am.pnu.com). The  
library was created from dissected whole female gonads.  
DNA sequencing was done by Washington University Genome  
Sequencing Center, St. Louis, MO."

BASE COUNT 99 a 46 c 88 g 56 t

ORIGIN

Alignment Scores:

Pred. No.:	1.81e+03	Length:	289
Score:	44.00	Matches:	6
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	91.67%	Indels:	0
DB:	14	Gaps:	0

US-09-877-606-7 (1-7) x BQ095463 (1-289)

OY 2 TrpAsnTrpPheAsnIle 7  
|||||

Db 232 TGGAACTGCTTCATATATA 249

SOURCE Gossypium arboreum.  
ORGANISM Gossypium arboreum  
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Malvales; Malvaceae; Gossypium.

REFERENCE 1 (bases 1 to 877)  
AUTHORS Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.  
TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber  
JOURNAL Unpublished (2000)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: twing@clemson.edu  
Seq primer: TAATACGACTCCTATGAGG  
High quality sequence stop: 680.

FEATURES  
SOURCE Location/Qualifiers  
1..877  
/organism="Gossypium arboreum"  
/strain="AKA"  
/cultivar="8400"  
/db\_xref="taxon:29729"  
/clone="GA\_EA0024024f"  
/clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
/tissue\_type="Fibers isolated from bolls harvested 7-10 dpa"  
/lab\_host="E. coli"  
/note="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"

BASE COUNT 201 a 183 c 231 g 261 t 1 others  
ORIGIN

Alignment Scores:  
Pred. No.: 4.34e+03 Length: 877  
Score: 46.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 85.71% Mismatches: 0  
Query Match: 95.83% Indels: 0  
Gaps: 0

US-09-877-606-7 (1-7) x BG444613 (1-877)

OY 1 LeutrpaSnrTpheaSnle 7  
DB 408 CTGTGGAAGTGTTCACCTT 428

RESULT 4  
AV809810 433 bp mRNA linear EST 29-MAR-2002  
LOCUS AV809810 Arabidopsis thaliana cDNA clone RAFL09-61-F20 3,  
DEFINITION mRNA sequence.  
ACCESSION AV809810  
VERSION AV809810.1 GI:19843795  
KEYWORDS EST.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 433)  
AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arikawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
TITLE Large scale analysis of Arabidopsis full-length cDNA (2002p)  
JOURNAL Unpublished (2002)  
COMMENT Contact: Motoaki Seki  
Plant Functional Genomics Research Group  
RIKEN Genomic Sciences Center  
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

Tel: 81-298-36-4359  
Fax: 81-298-36-9060  
Email: mseki@rkc.riken.go.jp  
An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified Lambda FLX-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified pBluescript vector. Please visit our web site ([http://www.gsc.riken.go.jp/e/plant/index\\_e.html](http://www.gsc.riken.go.jp/e/plant/index_e.html)) for further details.

FEATURES  
SOURCE Location/Qualifiers  
1..433  
/organism="Arabidopsis thaliana"  
/db\_xref="taxon:3702"  
/clone="RAFL09-61-F20"  
/clone\_lib="RAFL9"  
/dev\_stage="plants at various developmental stages from germination to mature seeds"  
/lab\_host="DH10B"  
/note="Site\_1: BamHI; Site\_2: SalI; subjected to dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24 hr) treatments"

BASE COUNT 183 a 81 c 61 g 108 t  
ORIGIN

Alignment Scores:  
Pred. No.: 2.27e+03 Length: 433  
Score: 45.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 85.71% Mismatches: 0  
Query Match: 93.75% Indels: 0  
Gaps: 0

US-09-877-606-7 (1-7) x AV809810 (1-433)

OY 1 LeutrpaSnrTpheaSnle 7  
DB 357 CTGTGGAAGTGTATATATT 337

RESULT 5  
B28259 544 bp DNA linear GSS 13-OCT-1997  
LOCUS B28259 Arabidopsis thaliana genomic clone T7N18, DNA  
DEFINITION T7N18TF TAWC Arabidopsis thaliana genomic clone T7N18, DNA  
ACCESSION B28259  
VERSION B28259.1 GI:2514225  
KEYWORDS GSS.  
SOURCE thale cress.  
ORGANISM Arabidopsis thaliana  
Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 544)  
AUTHORS Rounsley, S.D., Kelley, J.M., Field, C.E., Craven, M.B., Adams, M.D. and Venter, J.C.  
TITLE Use of a BAC End Sequence Database to Identify Minimal Overlaps for Arabidopsis genomic Sequencing  
JOURNAL Unpublished (1997)  
COMMENT Other-GSSs: T7N18TRC  
Contact: Steve Rounsley  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0208  
Fax: 301 838 0208  
Email: rounsley@igf.org  
Seq primer: M13-21  
Class: BAC ends  
High quality sequence stop: 544.

FEATURES  
SOURCE Location/Qualifiers  
1..544  
/organism="Arabidopsis thaliana"  
/strain="Columbia"



TITLE 'M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausern, A. and Wright, D., Weiss, R.  
Mouse whole genome scaffolding with paired end reads from 10kb  
plasmid inserts

JOURNAL Unpublished (2000)

COMMENT Contact: Robert B. Weiss  
University of Utah Genome Center  
University of Utah  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: ddunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0180 row: G column: 10  
Seq primer: CACACAGAAACAGCTATGACC  
Class: plasmid ends  
High quality sequence stop: 664.

## FEATURES

source

1. 664  
/organism="Mus musculus"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="U06C2M0180G10"  
/clone\_id="Mouse 10kb plasmid U06C1M library"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/note="Vector: PMD42nv: Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PMD42 (g1473114|g1473114|g1473114|g1473114), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 206 a 154 c 152 g 150 t 2 others

ORIGIN

## Alignment Scores:

Pred. No.: 2.1e+03 Length: 664  
Score: 47.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 85.71% Mismatches: 0  
Query Match: 97.92% Indels: 0  
DB: 17 Gaps: 0

US-09-877-606-7 (1-7) x AZ868788 (1-664)

Oy 1 Leutrpastrphea5nile 7  
|||||

Db 201 CTGTGGAAGTGTTCATATGTT 221

RESULT 2  
AW733876/c 551 bp mRNA linear EST 03-DEC-2001  
LOCUS AW733876  
DEFINITION SK78811.y1 Gm-cl016 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:  
ACCESSION AW733876  
VERSION AW733876.1 GI:7639556  
KEYWORDS EST.  
SOURCE soybean.  
ORGANISM Glycine max

REFERENCE Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.

AUTHORS 1 (bases 1 to 551)  
Shoemaker, R., Keim, P., Vockin, L., Erpelting, J., Corryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

TITLE Public Soybean EST Project  
JOURNAL Unpublished (1999)  
COMMENT Contact: Shoemaker R/Public Soybean EST Project  
Public Soybean EST Project  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu

## FEATURES

source

1. 551  
/organism="Glycine max"  
/db\_xref="taxon:3847"  
/clone="GENOME SYSTEMS CLONE ID: Gm-cl016-10245"  
/clone\_id="Gm-cl016"  
/tissue-type="Immature flowers of field grown plants"  
/lab\_host="XL10-Gold"  
/note="Vector: pBluescript II XR: Site 1: EcoRI; Site 2: XhoI. This cDNA library was constructed from mRNA isolated from immature flowers of field grown plants. The cDNA library was prepared using the Stratagene pBluescript II XR library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly (dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into XL10-Gold host cells. This library was constructed by Dr. Randy Shoemaker and Dr. John Erpelting."

BASE COUNT 188 a 136 c 62 g 165 t

ORIGIN

## Alignment Scores:

Pred. No.: 2.27e+03 Length: 551  
Score: 46.00 Matches: 6  
Percent Similarity: 100.00% Conservative: 1  
Best Local Similarity: 85.71% Mismatches: 0  
Query Match: 95.83% Indels: 0  
DB: 10 Gaps: 0

US-09-877-606-7 (1-7) x AW733876 (1-551)

Oy 1 Leutrpastrphea5nile 7  
|||||

Db 60 CTTTGAATGTGTTAACTG 40

RESULT 3  
BG444613 877 bp mRNA linear EST 15-MAR-2001  
LOCUS BG444613  
DEFINITION GA\_Ea0024024f Gossypium arboreum 7-10 dpa fiber library Gossypium  
arboresum cDNA clone GA\_Ea0024024f, mRNA sequence.  
ACCESSION BG444613  
VERSION BG444613.1 GI:13354265  
KEYWORDS EST.

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:52:44 ; Search time 197.31 Seconds  
(without alignments)  
574.569 Million cell updates/sec

Title: us-09-877-606-7  
Perfect score: 48  
Sequence: 1 LMNWFNI 7

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n model -DEV=xlh  
-O=/cgn2.1/USPTO.spool/US09877606/runat.18032003.134800.17705/app.query.fasta.1.1059  
-DB=EST -QFMT=fastap -SURFIX=est -MINMATCH=0 -DOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -EMD=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOCALLGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09877606.@cgn2.1.4134.@runat.18032003.134800.17705 -NCPU=6 -ICPU=3  
-NO.XLPHY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-MARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST: \*  
1: em\_estba: \*  
2: em\_esthum: \*  
3: em\_estlu: \*  
4: em\_estnu: \*  
5: em\_estov: \*  
6: em\_estpl: \*  
7: em\_estro: \*  
8: em\_hic: \*  
9: gb\_est1: \*  
10: gb\_est2: \*  
11: gb\_hic: \*  
12: gb\_est3: \*  
13: gb\_est4: \*  
14: gb\_est5: \*  
15: em\_estfun: \*  
16: em\_estom: \*  
17: gb\_gss: \*  
18: em\_gss\_hum: \*  
19: em\_gss\_luv: \*  
20: em\_gss\_pin: \*  
21: em\_gss\_vit: \*  
22: em\_gss\_fun: \*  
23: em\_gss\_mam: \*  
24: em\_gss\_mus: \*  
25: em\_gss\_other: \*  
26: em\_gss\_pro: \*  
27: em\_gss\_tod: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	47	97.9	664	17	AZ868788
2	46	95.8	551	10	AW733876
3	46	95.8	877	12	BG444613
4	45	93.8	433	10	AV809810
5	45	93.8	544	17	B28259
6	45	93.8	647	17	B67686
7	44	91.7	289	14	B0095463
8	44	91.7	327	9	A1938434
9	44	91.7	376	12	BG604081
10	44	91.7	409	10	AV799546
11	44	91.7	418	10	AW306467
12	44	91.7	429	10	AW395321
13	44	91.7	436	12	BF324158
14	44	91.7	437	10	AW156192
15	44	91.7	437	10	AW458604
16	44	91.7	440	10	AV785668
17	44	91.7	451	10	AW459039
18	44	91.7	458	14	BQ252572
19	44	91.7	459	10	AW396020
20	44	91.7	462	10	BE059144
21	44	91.7	466	14	BQ081613
22	44	91.7	475	12	BE806193
23	44	91.7	475	13	BJ388102
24	44	91.7	479	12	BG404967
25	44	91.7	481	10	AW306398
26	44	91.7	481	10	AW396674
27	44	91.7	487	12	BE805026
28	44	91.7	495	9	A1938282
29	44	91.7	495	12	BE805256
30	44	91.7	501	17	AO135880
31	44	91.7	509	14	BQ627823
32	44	91.7	514	13	BM528336
33	44	91.7	528	13	BJ390624
34	44	91.7	528	14	BQ297200
35	44	91.7	534	12	BG045640
36	44	91.7	538	9	AU199833
37	44	91.7	542	14	BQ299106
38	44	91.7	545	14	BM732434
39	44	91.7	551	13	BM527538
40	44	91.7	552	13	BM528410
41	44	91.7	561	14	BM732219
42	44	91.7	583	13	BJ413453
43	44	91.7	593	13	BI165630
44	44	91.7	593	17	AQ249789
45	44	91.7	616	9	AU204656

ALIGNMENTS

RESULT 1  
AZ868788  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS

AZ868788 664 bp DNA linear GSS 21-FEB-2001  
clone UUGC2M0180G10 R. DNA sequence.  
AZ868788  
AZ868788.1 GI:13072452  
GSS.  
house mouse.  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 664)  
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly

Thu Mar 27 11:38:01 2003

us-09-877-606-7.rspt

Page 6

Search completed: March 18, 2003, 14:36:06  
Job time : 4.93103 secs

---

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CAM70;  
RX MEDLINE=20414627; PubMed=10957729;  
RA Peter F.N., Etel M.N., Judith T.N., George A.A., Laura E.T.,  
RT John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.;  
RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral  
RT Variants in Cameroon."  
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CAM70;  
RA Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemni G.A., Eno L.T.,  
RA Nengasong J.N., Gao F., Rayfield M., Folks T.M., Pleniszek D.,  
RA Lal R.B.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF252128; AAC14337.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KM Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 118  
SQ SEQUENCE 118 AA; 14346 MW; CE454BE980AEFB1B CRC64;  
Query Match 100.0%; Score 48; DB 15; Length 118;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LMNFMNI 7  
Db 105 LMNFMNI 111

RESULT 13  
Q8UOX7  
ID Q8UOX7 PRELIMINARY; PRT; 121 AA.  
AC Q8UOX7;  
DT 01-MAR-2002 (TREMBlrel. 20, Created)  
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-982M072M;  
RX MEDLINE=21602569; PubMed=11739704;  
RA Trask S.A., Derdeyn C.A., Fidel U., Chen Y., Meleth S., Kasolo F.,  
RA Misonda R., Hunter E., Gao F., Allen S., Hahn B.H.;  
RT "Molecular epidemiology of human immunodeficiency virus type 1  
RT transmission in a heterosexual cohort of discordant couples in  
RT Zambia."  
RL J. Virol. 76:397-405(2002).  
DR EMBL: AF405173; AA66691.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
FT NON\_TER 1  
FT NON\_TER 121  
SQ SEQUENCE 121 AA; 14684 MW; F3AC3DB2FBE8802 CRC64;  
Query Match 100.0%; Score 48; DB 15; Length 121;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LMNFMNI 7  
Db 105 LMNFMNI 111

RESULT 14  
Q9YXR6  
ID Q9YXR6 PRELIMINARY; PRT; 122 AA.  
AC Q9YXR6;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Envelope glycoprotein immunodominant region (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RJ96BRP004;  
RA Tanuri A., Swanson P.A., Devare S.G., Berro O.J., Savedra A.,  
RA Costa L.J., Telles J.G., Brindello R., Schable C., Pleniszek D.,  
RA Rayfield M.;  
RL "HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil."  
RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF034037; AAC79289.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KM Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 122  
SQ SEQUENCE 122 AA; 14702 MW; 873C9A28AE14BA78 CRC64;  
Query Match 100.0%; Score 48; DB 15; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LMNFMNI 7  
Db 109 LMNFMNI 115

RESULT 15  
Q9YXP1  
ID Q9YXP1 PRELIMINARY; PRT; 122 AA.  
AC Q9YXP1;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Envelope glycoprotein immunodominant region (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RJ96BRP081;  
RA Tanuri A., Swanson P.A., Devare S.G., Berro O.J., Savedra A.,  
RA Costa L.J., Telles J.G., Brindello R., Schable C., Pleniszek D.,  
RA Rayfield M.;  
RL "HIV-1 subtypes among blood donors from Rio de Janeiro, Brazil."  
RT Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF034062; AAC79314.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KM Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 122  
SQ SEQUENCE 122 AA; 14778 MW; FD319ICE79ABB19D CRC64;  
Query Match 100.0%; Score 48; DB 15; Length 122;  
Best Local Similarity 100.0%; Pred. No. 1.9;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 LMNFMNI 7  
Db 109 LMNFMNI 115



FT NON\_TER 118 118  
SQ SEQUENCE 118 AA; 14257 MW; C11CFF6BDEFCBC50 CRC64;  
Query Match  
Best Local Similarity 100.0%; Score 48; DB 15; Length 118;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7  
Db 105 LMNWFNI 111

RESULT 6  
Q9E5R9 PRELIMINARY; PRT; 118 AA.  
AC Q9E5R9;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CAM16;  
RX MEDLINE=20414627; PubMed=10957729;  
RA Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T.,  
RA John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.;  
RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral  
RT Variants in Cameroon."  
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CAM16;  
RA Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemuji G.A., Eno L.T.,  
RA Ngenasong J.N., Gao F., Rayfield M., Folks T.M., Plenlazez D.,  
RA Lal R.B.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF252103; AAG14311.1; -;  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 118 118  
SQ SEQUENCE 118 AA; 14159 MW; 54739AADD14A51F7 CRC64;  
Query Match  
Best Local Similarity 100.0%; Score 48; DB 15; Length 118;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7  
Db 105 LMNWFNI 111

RESULT 7  
Q9E5R8 PRELIMINARY; PRT; 118 AA.  
AC Q9E5R8;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CAM161;  
RX MEDLINE=20414627; PubMed=10957729;

RA Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T.,  
RA John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.;  
RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral  
RT Variants in Cameroon."  
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CAM161;  
RA Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemuji G.A., Eno L.T.,  
RA Ngenasong J.N., Gao F., Rayfield M., Folks T.M., Plenlazez D.,  
RA Lal R.B.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF252103; AAG14312.1; -;  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 118 118  
SQ SEQUENCE 118 AA; 14244 MW; CBE698BE4A3748C3 CRC64;  
Query Match  
Best Local Similarity 100.0%; Score 48; DB 15; Length 118;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7  
Db 105 LMNWFNI 111

RESULT 8  
Q9E5Q8 PRELIMINARY; PRT; 118 AA.  
AC Q9E5Q8;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CAM203;  
RX MEDLINE=20414627; PubMed=10957729;  
RA Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T.,  
RA John N.N., Feng G., Mark R., Thomas F.M., Danuta P., Renu L.B.;  
RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral  
RT Variants in Cameroon."  
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CAM203;  
RA Fonjongo P.N., Mpoudi E.N., Torimiro J.N., Alemuji G.A., Eno L.T.,  
RA Ngenasong J.N., Gao F., Rayfield M., Folks T.M., Plenlazez D.,  
RA Lal R.B.;  
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF252113; AAG14322.1; -;  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1 1  
FT NON\_TER 118 118  
SQ SEQUENCE 118 AA; 14173 MW; 488033169BCD3CB6 CRC64;  
Query Match  
Best Local Similarity 100.0%; Score 48; DB 15; Length 118;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMNWFNI 7  
Db 105 LMNWFNI 111

Q9YRT3 PRELIMINARY; PRT; 117 AA.  
ID Q9YRT3  
AC Q9YRT3;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=96GUSM11;  
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,  
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,  
RA Jones J., Ward J.;  
RT "Surveillance of Central African Nationals living in the United States  
RT reveals multiple subtypes of HIV-1 Group M and Group O.";  
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF096332; AAD04407.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 117  
SQ SEQUENCE 117 AA; 14088 MW; 32ADA48B31BAFF4 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMWNI 7  
Db 109 LMWNI 115

RESULT 3  
Q9YRS7 PRELIMINARY; PRT; 117 AA.  
ID Q9YRS7  
AC Q9YRS7;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=96GUSM07;  
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,  
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,  
RA Jones J., Ward J.;  
RT "Surveillance of Central African Nationals living in the United States  
RT reveals multiple subtypes of HIV-1 Group M and Group O.";  
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF096339; AAD04414.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 117  
SQ SEQUENCE 117 AA; 14133 MW; ECBABE6A1CDA17D CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMWNI 7  
Db 109 LMWNI 115

RESULT 4  
Q9YR6 PRELIMINARY; PRT; 117 AA.  
ID Q9YR6  
AC Q9YR6;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=96GUSM8;  
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,  
RA Robbins K., Lal R., Storck C., Schable C.A., Wise H., Tetteh C.,  
RA Jones J., Ward J.;  
RT "Surveillance of Central African Nationals living in the United States  
RT reveals multiple subtypes of HIV-1 Group M and Group O.";  
RL Submitted (OCT-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF096350; AAD04425.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 117  
SQ SEQUENCE 117 AA; 13986 MW; 39B9C97DA14BF47E CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LMWNI 7  
Db 109 LMWNI 115

RESULT 5  
Q9ES6 PRELIMINARY; PRT; 118 AA.  
ID Q9ES6  
AC Q9ES6;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CAM103;  
RX MEDLINE=20414627; PubMed=10957729;  
RA Peter F.N., Eitel M.N., Judith T.N., George A.A., Laura E.T.,  
RA John N.N., Peng G., Mark R., Thomas F.M., Danuta P., Renu L.B.;  
RT "Presence of Diverse Human Immunodeficiency Virus Type 1 Viral  
RT Variants in Cameroon.";  
RL AIDS Res. Hum. Retroviruses 16:1319-1324(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CAM103;  
RA Fonjuno P.N., Mpoudi E.N., Torimiro J.N., Alemni G.A., Eno L.T.,  
RA Nengasong J.N., Gao F., Rayfield M., Folks T.M., Pieniazek D.,  
RA Lal R.B.;  
RL Submitted (APR-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF252095; AAC14304.1; -  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam; PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER 1

GenCore version 5.1.4-p5-4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:01 ; Search time 3.93103 Seconds  
(without alignments)  
366.908 Million cell updates/sec

Title: US-09-877-606-7  
Perfect score: 48  
Sequence: 1 LMNMFNI 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

- 1: SPREMBL\_21:\*
- 2: sp\_archaea:\*
- 3: sp\_bacteria:\*
- 4: sp\_fungi:\*
- 5: sp\_human:\*
- 6: sp\_invertebrate:\*
- 7: sp\_mammal:\*
- 8: sp\_mhc:\*
- 9: sp\_organelle:\*
- 10: sp\_phage:\*
- 11: sp\_plant:\*
- 12: sp\_rodent:\*
- 13: sp\_virus:\*
- 14: sp\_vertebrate:\*
- 15: sp\_unclassified:\*
- 16: sp\_virus:\*
- 17: sp\_bacteriophage:\*
- 18: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	48	100.0	117	15	O9YRT8	O9YRT8 human immun
2	48	100.0	117	15	O9YRT3	O9YRT3 human immun
3	48	100.0	117	15	O9YRS7	O9YRS7 human immun
4	48	100.0	117	15	O9YRK6	O9YRK6 human immun
5	48	100.0	118	15	O9ES56	O9ES56 human immun
6	48	100.0	118	15	O9ESR9	O9ESR9 human immun
7	48	100.0	118	15	O9ESR8	O9ESR8 human immun
8	48	100.0	118	15	O9ES08	O9ES08 human immun
9	48	100.0	118	15	O9ESQ2	O9ESQ2 human immun
10	48	100.0	118	15	O9ESP9	O9ESP9 human immun
11	48	100.0	118	15	O9ESP4	O9ESP4 human immun
12	48	100.0	118	15	O9ESP3	O9ESP3 human immun
13	48	100.0	121	15	O9BQX7	O9BQX7 human immun
14	48	100.0	122	15	O9YXR6	O9YXR6 human immun
15	48	100.0	122	15	O9YXP1	O9YXP1 human immun
16	48	100.0	122	15	O9YXN8	O9YXN8 human immun

17	48	100.0	122	15	O9YXN7	O9YXN7 human immun
18	48	100.0	122	15	O9YXN5	O9YXN5 human immun
19	48	100.0	122	15	O9YXN4	O9YXN4 human immun
20	48	100.0	122	15	O9YXN1	O9YXN1 human immun
21	48	100.0	122	15	O9YXN8	O9YXN8 human immun
22	48	100.0	122	15	O9YR03	O9YR03 human immun
23	48	100.0	122	15	O9Q1W9	O9Q1W9 human immun
24	48	100.0	122	15	O9Q1W6	O9Q1W6 human immun
25	48	100.0	122	15	O9Q1W0	O9Q1W0 human immun
26	48	100.0	122	15	O9Q1U9	O9Q1U9 human immun
27	48	100.0	122	15	O9EAA5	O9EAA5 human immun
28	48	100.0	122	15	O9EAA0	O9EAA0 human immun
29	48	100.0	122	15	O9EAA6	O9EAA6 human immun
30	48	100.0	122	15	O9ILK7	O9ILK7 human immun
31	48	100.0	122	15	O9ILK1	O9ILK1 human immun
32	48	100.0	122	15	O9ILK0	O9ILK0 human immun
33	48	100.0	122	15	O9ILQ7	O9ILQ7 human immun
34	48	100.0	122	15	O9ILQ1	O9ILQ1 human immun
35	48	100.0	122	15	O9ILP9	O9ILP9 human immun
36	48	100.0	122	15	O9ILP7	O9ILP7 human immun
37	48	100.0	122	15	O9ILP2	O9ILP2 human immun
38	48	100.0	122	15	O9ILJ1	O9ILJ1 human immun
39	48	100.0	122	15	O9OPY4	O9OPY4 human immun
40	48	100.0	122	15	O9ODP0	O9ODP0 human immun
41	48	100.0	122	15	O9ODS9	O9ODS9 human immun
42	48	100.0	122	15	O9ODP1	O9ODP1 human immun
43	48	100.0	122	15	O9ODK1	O9ODK1 human immun
44	48	100.0	126	15	O9YXY3	O9YXY3 human immun
45	48	100.0	127	15	O9YXK6	O9YXK6 human immun

ALIGNMENTS

RESULT 1

ID O9YRT8 PRELIMINARY; PRT: 117 AA.  
AC O9YRT8;  
DT 01-MAY-1999 (TREMREL. 10, Created)  
DT 01-MAY-1999 (TREMREL. 10, Last sequence update)  
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=96USNG31.  
RA Sullivan P.J., Do A., Ellenberger D.L., Pau C.-P., Paul S., Kalish M.,  
RA Robbins K., Lal R., Storch C., Schable C.A., Wise H., Tetteh C.,  
RA Jones J., Ward J.;  
RT "Surveillance of Central African Nationals living in the United States  
RT reveals multiple subtypes of HIV-1 Group M and Group O.";  
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF096327; AAD04402.1; -;  
DR InterPro: IPR000328; Env\_GP41.  
DR Pfam: PF00517; GP41; 1.  
KW Transmembrane.  
FT NON\_TER  
FT NON\_TER  
SQ SEQUENCE 117 AA; 14156 MW; F3C6C7324386A240 CRC64;

Query Match 100.0%; Score 48; DB 15; Length 117;  
Best Local Similarity 100.0%; Pred. No. 1.8; Indels 0; Gaps 0;  
Matches 7; Conservative 0; Mismatches 0;

OY 1 LMNMFNI 7  
Db 109 LMNMFNI 115

RESULT 2



**THIS PAGE BLANK (USPTO)**

FT	CARBOHYD	350	350	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	356	356	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	380	380	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	386	386	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	390	390	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	400	400	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	438	438	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	450	450	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	602	602	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	607	607	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	616	616	N-LINKED (GLCNAC. . .)	(POTENTIAL).
FT	CARBOHYD	628	628	N-LINKED (GLCNAC. . .)	(POTENTIAL).
SO	SEQUENCE	847 AA;	96466 MM;	CDIE33D/3AA5BCAE	CRC64; (POTENTIAL).

Query Match 89.6%; Score 43; DB 1; Length 847;  
Best Local Similarity 85.7%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMENI 7  
||||:|  
Db 660 LMNMFDI 666

Search completed: March 18, 2003, 14:34:03  
Job time : 2 secs

DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (SF162 isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 ON NCB1 TaxID=11691;  
 RX MEDLINE=90347835; PubMed=2384920;  
 RA "Cheng-Mayer C., Outroga M., Tung J.W., Dina D., Levy J.;  
 RT "Viral determinants of human immunodeficiency virus type 1 T-cell or  
 RL macropage tropism, cytopathogenicity, and CD4 antigen modulation.";  
 CC J. Virol. 64:4390-4398(1990).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M65024; AAA45072.1; -  
 DR HIV; M38428; ENV5SF162.  
 DR InterPro: IPR000328; Env-GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 502 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 847 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 118 73 BY SIMILARITY.  
 FT DISULFID 125 203 BY SIMILARITY.  
 FT DISULFID 130 194 BY SIMILARITY.  
 FT DISULFID 130 155 BY SIMILARITY.  
 FT DISULFID 126 245 BY SIMILARITY.  
 FT DISULFID 226 237 BY SIMILARITY.  
 FT DISULFID 294 328 BY SIMILARITY.  
 FT DISULFID 374 435 BY SIMILARITY.  
 FT DISULFID 381 408 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 195 195 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 260 260 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 274 274 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 293 293 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 352 352 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 454 454 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 628 628 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 847 AA; 96135 MW; 0A901317FD/FF2AB CRC64;

DB 660 LMNMEDI 666  
 RESULT 15  
 ID ENV\_HV1W2 STANDARD; PRT; 847 AA.  
 AC P05880;  
 DT 01-NOV-1988 (Rel. 09, Created)  
 DT 01-NOV-1988 (Rel. 09, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Envelope polypeptide gp160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (WMJ2 isolate) (HIV-1).  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 ON NCB1 TaxID=11705;  
 RX MEDLINE=86235450; PubMed=3012778;  
 RA Hahn B.H., Shaw G.M., Taylor M.E., Redfield R.R., Markham P.D.,  
 RA Salahuddin S.Z., Wong-Staal F., Gallo R.C., Parks E.S., Parks W.P.;  
 RT "Genetic variation in HTLV-III/LAV over time in patients with AIDS or  
 RT at risk for AIDS.";  
 RL Science 232:1548-1553(1986).  
 CC -I- MISCELLANEOUS: ISOLATES WMJ1, WMJ2, AND WMJ3 WERE OBTAINED FROM  
 CC BLOOD SAMPLES SPONTANEOUSLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO  
 CC WAS PERINATALLY INFECTED BY HER MOTHER.  
 CC -----  
 DR EMBL: M12507; AAB12990.1; -  
 DR HIV; M12507; ENV5WMJ2.  
 DR InterPro: IPR000328; Env-GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDS; Coat protein; Polypotein; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 29  
 FT CHAIN 30 501 EXTERIOR MEMBRANE GLYCOPROTEIN.  
 FT DISULFID 53 847 TRANSMEMBRANE GLYCOPROTEIN.  
 FT DISULFID 118 73 BY SIMILARITY.  
 FT DISULFID 125 202 BY SIMILARITY.  
 FT DISULFID 130 193 BY SIMILARITY.  
 FT DISULFID 130 152 BY SIMILARITY.  
 FT DISULFID 215 244 BY SIMILARITY.  
 FT DISULFID 225 236 BY SIMILARITY.  
 FT DISULFID 293 326 BY SIMILARITY.  
 FT DISULFID 372 435 BY SIMILARITY.  
 FT DISULFID 379 408 BY SIMILARITY.  
 FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 134 134 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 183 183 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 259 259 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 89.6%; Score 43; DB 1; Length 847;  
 Best Local Similarity 85.7%; Pred. No. 19;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMEDI 7

```

FT CHAIN 490 843 TRANSMEMBRANE GLYCOPROTEIN.
FT TRANSEM 738 755 POTENTIAL.
FT DISULFID 53 73 BY SIMILARITY.
FT DISULFID 118 201 BY SIMILARITY.
FT DISULFID 125 192 BY SIMILARITY.
FT DISULFID 130 155 BY SIMILARITY.
FT DISULFID 214 243 BY SIMILARITY.
FT DISULFID 224 235 BY SIMILARITY.
FT DISULFID 292 326 BY SIMILARITY.
FT DISULFID 373 432 BY SIMILARITY.
FT DISULFID 380 405 BY SIMILARITY.
FT CARBOHYD 87 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 193 193 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 258 258 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 285 285 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 389 389 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 435 435 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 450 450 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 598 598 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 803 803 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 843 AA; 95648 MW; C69DED971C918B71 CRC64;

Query Match 89.6%; Score 43; DB 1; Length 843;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7
Db 656 LMNMFNI 662

RESULT 13
ENV_HV1ND STANDARD; PRT; 846 AA.
AC P18799;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide gp160 precursor [contains: exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (NDK isolate) (HIV-1).
OC Viruses; Retroviridae; Retroviridae; Lentivirus.
OX NCBI_TaxID=11695;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90034200; PubMed=2806917;
RA Spire B., Sire J., Zachar V., Rey F., Barre-Sinoussi F., Galibert F.,
Hampé A., Chermann J.C.;
RT "Nucleotide sequence of HIV1-NDK: a highly cytopathic strain of the
human immunodeficiency virus."
RL Gene 81:275-284(1989).
CC -1- MISCELLANEOUS: NDK, ISOLATED FROM A ZAIRIAN PATIENT AFFECTED WITH
AIDS, AND IS A HIGHLY CYTOPATHOGENIC STRAIN.

```

```

CC CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M27323; AAA44873.1; -.
DR PIR; J00066; VCLND.
DR HIV; M27323; ENV5NDK.
DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
FT CHAIN 1 29
FT CHAIN 30 501
FT CHAIN 502 846
FT DISULFID 53 73 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT DISULFID 118 200 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFID 125 191 BY SIMILARITY.
FT DISULFID 130 152 BY SIMILARITY.
FT DISULFID 213 242 BY SIMILARITY.
FT DISULFID 223 234 BY SIMILARITY.
FT DISULFID 291 328 BY SIMILARITY.
FT DISULFID 374 435 BY SIMILARITY.
FT DISULFID 381 408 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 182 182 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 229 229 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 257 257 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 284 284 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 351 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 388 388 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 438 438 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 451 451 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 452 452 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 601 601 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 627 627 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 846 AA; 96476 MW; 8A3B9DA527DE2E83 CRC64;

Query Match 89.6%; Score 43; DB 1; Length 846;
Best Local Similarity 85.7%; Pred. No. 19;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7
Db 659 LMNMFNI 665

RESULT 14
ENV_HV1SI STANDARD; PRT; 847 AA.
AC P19550;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide gp160 precursor [contains: exterior membrane

```





```

FT CARBOHD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FED CRC64;

Query Match 100.0%; Score 48; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LMNMENI 7
Db 669 LMNMENI 675

RESULT 8
ENV_HV1H3 STANDARD; PRT: 856 AA.
ID ENV_HV1H3
AC P04624;
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11707;
RN [1]
RP MEDLINE=85228248; Pubmed=2988795;
RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
Shaw G.M., Wong-Staal F., Reddy E.P.;
RT "HIV-III env gene products synthesized in E. coli are recognized by
antibodies present in the sera of AIDS patients.";
RL Cell 41:979-986(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL_M14100; AAA44679.1; -.
DR HIV; M14100; ENVSHXB3.
DR InterPro: IPR000328; ENV_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
Signal.
FT CHAIN 1 511 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT SIGNAL 30
FT CHAIN 512 856 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFD 119 205 BY SIMILARITY.
FT DISULFD 126 196 BY SIMILARITY.
FT DISULFD 131 157 BY SIMILARITY.
FT DISULFD 131 157 BY SIMILARITY.
FT DISULFD 218 247 BY SIMILARITY.
FT DISULFD 228 239 BY SIMILARITY.
FT DISULFD 296 331 BY SIMILARITY.
FT DISULFD 378 445 BY SIMILARITY.

```

```

FT DISULFD 385 418 BY SIMILARITY.
FT CARBOHD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97188 MW; 3373C68BB84C1AFC CRC64;

Query Match 100.0%; Score 48; DB 1; Length 856;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LMNMENI 7
Db 669 LMNMENI 675

RESULT 9
ENV_HV1LM STANDARD; PRT: 856 AA.
ID ENV_HV1LM
AC Q70626;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (LM12.3 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=82834;
RN [1]
RP MEDLINE=95127297; Pubmed=7826699;
RX Reitz M.S. Jr., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.M.,
Shaw G.M., Kong L.T., Weiss S.H., Waters D., Gallo R.C., Blattner W.;
RT "Viral variability and serum antibody response in a laboratory worker
infected with HIV type 1 (HIV type IIB).";
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----

```

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M15654; AAAA4205.1; -.  
 DR PIR: A03973; VCLJH3.  
 DR HIV; M15654; ENV5BH102.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDs; Coat protein; Polypeptide; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 511  
 FT DISULFID 512 856  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 445  
 FT DISULFID 385 418  
 FT DISULFID 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 234 234  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 289 289  
 FT CARBOHYD 295 295  
 FT CARBOHYD 301 301  
 FT CARBOHYD 332 332  
 FT CARBOHYD 339 339  
 FT CARBOHYD 356 356  
 FT CARBOHYD 366 366  
 FT CARBOHYD 386 386  
 FT CARBOHYD 392 392  
 FT CARBOHYD 397 397  
 FT CARBOHYD 406 406  
 FT CARBOHYD 448 448  
 FT CARBOHYD 463 463  
 FT CARBOHYD 611 611  
 FT CARBOHYD 616 616  
 FT CARBOHYD 625 625  
 FT CARBOHYD 637 637  
 FT CARBOHYD 674 674  
 FT CARBOHYD 750 750  
 FT CARBOHYD 816 816  
 SQ SEQUENCE 856 AA; 97224 MW; 0BFB1A18931BB27 CRC64;  
 Query Mch 100.0%; Score 48; DB 1; Length 856;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ENV\_HV1H2  
 ID ENV\_HV1H2 STANDARD; PRT; 856 AA.  
 AC P04578; 009779;  
 DT 13-AUG-1987 (Rel. 05, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Envelope glycoprotein gp160 precursor [Contains: Exterior membrane  
 DE glycoprotein (GP120); Transmembrane glycoprotein (GP41)].  
 GN ENV.  
 OS Human immunodeficiency virus type 1 (HXB2 isolate) (HIV-1).  
 OC Viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11706;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87299196; PubMed=3040055;  
 RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,  
 RA Gallo R.C., Wong-Staal F.;  
 RT "Complete nucleotide sequences of functional clones of the AIDS  
 RT virus.";  
 RL AIDS Res. Hum. Retroviruses 3:57-69(1987).  
 RN [2].  
 RP REVISIONS.  
 RA Ratner L., Fisher A., Jagodzinski L.L., Mitsuya H., Liou R.-S.,  
 RA Gallo R.C., Wong-Staal F.;  
 RL Submitted (APR-1997) to the EMBL/GenBank/DBJ databases.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: K03455; AAB50262.1; -.  
 DR EMBL: AF038399; AAB99976.1; -.  
 DR EMBL: AF038319; AAC82596.1; -.  
 DR HIV; K03455; ENV5SHB2.  
 DR InterPro: IPR000328; Env\_GP41.  
 DR InterPro: IPR000777; GP120.  
 DR Pfam: PF00516; GP120; 1.  
 DR Pfam: PF00517; GP41; 1.  
 KW AIDs; Coat protein; Polypeptide; Glycoprotein; Transmembrane;  
 KW Signal.  
 FT SIGNAL 1 30  
 FT CHAIN 31 511  
 FT DISULFID 512 856  
 FT DISULFID 54 74  
 FT DISULFID 119 205  
 FT DISULFID 126 196  
 FT DISULFID 131 157  
 FT DISULFID 218 247  
 FT DISULFID 228 239  
 FT DISULFID 296 331  
 FT DISULFID 378 445  
 FT DISULFID 385 418  
 FT DISULFID 88 88  
 FT CARBOHYD 136 136  
 FT CARBOHYD 141 141  
 FT CARBOHYD 156 156  
 FT CARBOHYD 160 160  
 FT CARBOHYD 186 186  
 FT CARBOHYD 197 197  
 FT CARBOHYD 230 230  
 FT CARBOHYD 241 241  
 FT CARBOHYD 262 262  
 FT CARBOHYD 276 276  
 FT CARBOHYD 289 289  
 FT CARBOHYD 295 295  
 FT CARBOHYD 301 301  
 FT CARBOHYD 332 332  
 FT CARBOHYD 339 339  
 FT CARBOHYD 356 356  
 FT CARBOHYD 366 366  
 FT CARBOHYD 386 386  
 FT CARBOHYD 392 392  
 FT CARBOHYD 397 397  
 FT CARBOHYD 406 406  
 FT CARBOHYD 448 448  
 FT CARBOHYD 463 463  
 FT CARBOHYD 611 611  
 FT CARBOHYD 616 616  
 FT CARBOHYD 625 625  
 FT CARBOHYD 637 637  
 FT CARBOHYD 674 674  
 FT CARBOHYD 750 750  
 FT CARBOHYD 816 816  
 SQ SEQUENCE 856 AA; 97224 MW; 0BFB1A18931BB27 CRC64;  
 Query Mch 100.0%; Score 48; DB 1; Length 856;  
 Best Local Similarity 100.0%; Pred. No. 3.5;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



```

FT CARBOHYD 462 462 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 608 608 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 613 613 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 634 634 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 853 AA; 97043 MW; 849B08CBAFF7008 CRC64;

Query Match
Best Local Similarity 100.0%; Score 48; DB 1; Length 853;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7
DB 666 LMNMFNI 672

ENV_HV126 STANDARD; PRT; 855 AA.
ID ENV_HV126
AC P04580;
DT 13-AUG-1987 (Rel. 05, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide precursor [Contains: Exterior membrane
glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].
GN ENV.
OS Human immunodeficiency virus type 1 (Zaire 6 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11708;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87248097; PubMed=3036660;
RA Srinivasan A., Anand R., Ranganathan P., Feorino P.,
RA Schochetman G., Curran J., Kalyanaraman V.S., Luciw P.A.,
RA Sanchez-Pescador R.;
RT "Molecular characterization of human immunodeficiency virus from
RT Zaire: nucleotide sequence analysis identifies conserved and variable
RT domains in the envelope gene.";
RL Gene 52:71-82(1987).

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: K03458; AAA45380.1; -.
DR PIR: D26192; VCLJ2R.
DR HIV: K03458; ENV526.
DR InterPro: IPR000328; Env_Gp41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120.1.
DR Pfam: PF00517; GP41.1.
KW AIDS; Coat protein; Glycoprotein; Transmembrane;
KW Signal.
FT SIGNAL 1 30
FT CHAIN 31 510 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 511 855 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFD 53 73 BY SIMILARITY.
FT DISULFD 118 207 BY SIMILARITY.
FT DISULFD 125 198 BY SIMILARITY.
FT DISULFD 130 155 BY SIMILARITY.
FT DISULFD 220 249 BY SIMILARITY.
FT DISULFD 230 241 BY SIMILARITY.
FT DISULFD 298 332 BY SIMILARITY.
FT DISULFD 378 444 BY SIMILARITY.
FT DISULFD 385 417 BY SIMILARITY.
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 140 140 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 189 189 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 243 243 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 264 264 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 297 297 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 610 610 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 636 636 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 673 673 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 855 AA; 96971 MW; 384D3D6E239C3457 CRC64;

Query Match
Best Local Similarity 100.0%; Score 48; DB 1; Length 855;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7
DB 668 LMNMFNI 674

ENV_HV1B1 STANDARD; PRT; 856 AA.
ID ENV_HV1B1
AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide precursor [Contains: Exterior membrane
glycoprotein (Gp120); Transmembrane glycoprotein (Gp41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85111123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Doran E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R., Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Ghrayeb J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RL Nature 313:277-284(1985).

[2]
DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
MEDLINE=90285159; PubMed=2355006;
RX Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
RT potential glycosylation sites of the type 1 recombinant human
RT immunodeficiency virus envelope glycoprotein (gp120) expressed in
RT Chinese hamster ovary cells.";
RL J. Biol. Chem. 265:10373-10382(1990).

```

RX	MEDLINE=90317877; PubMed=1695254;
RA	Stevenson M., Haggerty S., Lamonia C., Mann A.M., Meier C., Waslask A.;
RT	"Cloning and characterization of human immunodeficiency virus type 1 variants diminished in the ability to induce syncytium-independent cytolysis.";
RL	J. Virol. 64:3792-3803(1990).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ) or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .
CC	-----
DR	EMBL, M33943; AAA44850.1; "
DR	HIV, M33943; ENVSMPA.
DR	InterPro; IPR000328; Env_GP41.
DR	InterPro; IPR000777; GP120.
DR	Pfam; PF00516; GP120; 1.
DR	Pfam; PF00517; GP41; 1.
KW	AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.
KW	SIGNAL.
FT	CHAIN 1 30
FT	CHAIN 510 853
FT	DISELFD 54 74
FT	DISELFD 119 203
FT	DISELFD 126 194
FT	DISELFD 131 157
FT	DISELFD 131 157
FT	DISELFD 216 245
FT	DISELFD 226 237
FT	DISELFD 294 329
FT	DISELFD 376 443
FT	DISELFD 383 416
FT	CARBOHD 88 88
FT	CARBOHD 136 136
FT	CARBOHD 141 141
FT	CARBOHD 156 156
FT	CARBOHD 160 160
FT	CARBOHD 186 186
FT	CARBOHD 195 195
FT	CARBOHD 232 232
FT	CARBOHD 239 239
FT	CARBOHD 260 260
FT	CARBOHD 274 274
FT	CARBOHD 287 287
FT	CARBOHD 293 293
FT	CARBOHD 299 299
FT	CARBOHD 330 330
FT	CARBOHD 354 354
FT	CARBOHD 384 384
FT	CARBOHD 390 390
FT	CARBOHD 395 395
FT	CARBOHD 404 404
FT	CARBOHD 446 446
FT	CARBOHD 461 461
FT	CARBOHD 609 609
FT	CARBOHD 614 614
FT	CARBOHD 623 623
FT	CARBOHD 635 635
FT	CARBOHD 672 672
FT	CARBOHD 748 748
FT	CARBOHD 814 814
SO	SEQUENCE 853 AA; 96912 MW; 33778993B6F22ABA CRC64;
QY	Query Match 100.0%; Score 48; DB 1; Length 853; Best Local Similarity 100.0%; Pred. No. 3.5; Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0; 

ID	ENV_HV122	STANDARD;	PRT;	853 AA.
DB	667	LMNMENI	673	
RESULT 4				
ENV_HV122				
AC	ENV_HV122	STANDARD;	PRT;	853 AA.
AC	P12487;			
DT	01-OCT-1989 (Rel. 12, Created)			
DT	01-OCT-1989 (Rel. 12, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Envelope polyprotein GP160 precursor [contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].			
GN	ENV.			
OS	Human immunodeficiency virus type 1 (22/CDC-234 isolate) (HIV-1).			
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.			
OX	NCBI_TaxID=11683;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Theodore T., Buckler-White A.;			
RL	Submitted (NOV-1988) to the HIV data bank.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb.ch/announce/">http://www.isb.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb.slb.ch">license@isb.slb.ch</a> ).			
CC	-----			
DR	EMBL; M22639; AAA45370.1; -			
DR	HIV; M22639; ENV\$Z226.			
DR	InterPro; IPR000328; Env_GP41.			
DR	InterPro; IPR000777; GP120.			
DR	Pfam; PF00516; GP120; 1.			
DR	Pfam; PF00517; GP41; 1.			
KW	AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;			
KW	Signal.			
FT	CHAIN	1	31	BY SIMILARITY.
FT	CHAIN	32	508	EXTERIOR MEMBRANE GLYCOPROTEIN.
FT	DISULFID	53	853	TRANSMEMBRANE GLYCOPROTEIN.
FT	DISULFID	118	206	BY SIMILARITY.
FT	DISULFID	125	197	BY SIMILARITY.
FT	DISULFID	130	154	BY SIMILARITY.
FT	DISULFID	219	248	BY SIMILARITY.
FT	DISULFID	229	240	BY SIMILARITY.
FT	DISULFID	297	330	BY SIMILARITY.
FT	DISULFID	376	442	BY SIMILARITY.
FT	DISULFID	383	415	BY SIMILARITY.
FT	CARBOHYD	87		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	137		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	144		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	153		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	157		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	185		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	188		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	198		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	235		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	242		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	263		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	277		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	290		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	296		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	331		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	338		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	353		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	384		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	390		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	402		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	441		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	445		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	458		N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	459		N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857785F1 CRC64;

```

```

Query Match 100.0%; Score 48; DB 1; Length 851;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 LMNMFNI 7
Db 664 LMNMFNI 670

```

```

RESULT 2
ENV_HV1BN STANDARD; PRT; 852 AA.
ID ENV_HV1BN
AC P12488;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide gp160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DE ENV.
OS Human immunodeficiency virus type 1 (1BR isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11693;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89085613; PubMed=2789516;
RA Anand R., Thayer R., Srinivasan A., Nayyar S., Gardner M., Luciw P., Dandekar S.;
RT "Biological and molecular characterization of human immunodeficiency virus (HIV-1BR) from the brain of a patient with progressive dementia.";
RL Virology 168:79-89(1989).
CC -I MISCELLANEOUS: THIS VIRUS IS CYTOPATHICALLY ACTIVE AND WAS HARVESTED FROM THE BRAIN TISSUE OF A NEUROLOGICAL AIDS PATIENT.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

```

```

CC EMBL; M21098; AAA44221.1; -
CC PIR; A31667; VCLJBR.
CC HIV; M21098; ENVSBRVA.

```

```

DR InterPro; IPR000328; Env_GP41.
DR InterPro; IPR000777; GP120.
DR Pfam; PF00516; GP120; 1.
DR Pfam; PF00517; GP41; 1.
KW AIDS; Coat protein; Polypeptide; Glycoprotein; Transmembrane; Signal.

```

```

FT CHAIN 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 508 852 TRANSMEMBRANE GLYCOPROTEIN.
FT DISULFD 54 74 BY SIMILARITY.
FT DISULFD 119 205 BY SIMILARITY.
FT DISULFD 126 196 BY SIMILARITY.
FT DISULFD 131 155 BY SIMILARITY.
FT DISULFD 218 247 BY SIMILARITY.
FT DISULFD 228 239 BY SIMILARITY.
FT DISULFD 296 330 BY SIMILARITY.
FT DISULFD 376 439 BY SIMILARITY.
FT DISULFD 383 412 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 135 135 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 138 138 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 154 154 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 158 158 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 354 354 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 360 360 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 396 396 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 607 607 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 612 612 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 621 621 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 633 633 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 670 670 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 852 AA; 97203 MW; 2BB866345DEC915F CRC64;

```

```

Query Match 100.0%; Score 48; DB 1; Length 852;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 1 LMNMFNI 7
Db 665 LMNMFNI 671

```

```

RESULT 3
ENV_HV1MF STANDARD; PRT; 853 AA.
ID ENV_HV1MF
AC P19551;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide gp160 precursor [Contains: Exterior membrane glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
DE ENV.
OS Human immunodeficiency virus type 1 (MRA isolate) (HIV-1).
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11704;
RN [1]
RP SEQUENCE FROM N.A.

```



Query Match 89.6%; Score 43; DB 2; Length 357;  
Best Local Similarity 85.7%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||:|  
Db 170 LMNMFNI 176

## RESULT 13

S21996  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S70422; S21996  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70422  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-357 <STE2>  
A:Cross-references: EMBL:X61356; NID:960181; PIDN:CAA43624.1; PID:91067129  
A:Note: submitted to the EMBL Data Library, July 1991  
C:Superfamily: type E retrovirus env polyprotein

Query Match 89.6%; Score 43; DB 2; Length 357;  
Best Local Similarity 85.7%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||:|  
Db 170 LMNMFNI 176

## RESULT 14

S21998  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 28  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S21998; S70425  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi  
A:Reference number: S21990  
A:Accession: S21998  
A:Molecule type: DNA  
A:Residues: 1-358 <STE1>  
A:Cross-references: EMBL:X61359; NID:960182; PIDN:CAA43630.1; PID:960183  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebr  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70425  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-222, 'X', 224-358 <STE2>  
A:Cross-references: EMBL:X61359; NID:960182; PIDN:CAA43630.1; PID:960183  
C:Superfamily: type E retrovirus env polyprotein

Query Match 89.6%; Score 43; DB 2; Length 358;  
Best Local Similarity 85.7%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||:|  
Db 171 LMNMFNI 177

## RESULT 15

S22002  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 3L  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S22002; S70418  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as dete  
A:Reference number: S21990  
A:Accession: S22002  
A:Molecule type: DNA  
A:Residues: 1-358 <STE1>  
A:Cross-references: EMBL:X61352; NID:960186; PIDN:CAA43616.1; PID:960187  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70418  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333, 'X', 335-358 <STE2>  
A:Cross-references: EMBL:X61352; NID:960186  
C:Superfamily: type E retrovirus env polyprotein

Query Match 89.6%; Score 43; DB 2; Length 358;  
Best Local Similarity 85.7%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||:|  
Db 171 LMNMFNI 177

Search completed: March 18, 2003, 14:37:20  
Job time : 4.24138 secs

F:517-861/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,146,161,165,191,202,239,246,267,300,306,337,344,361,391,397,402,411  
F:616,621,630,642,679,755,821/inding site: carbohydrate (Asn) (covalent) #status predic

Query Match 100.0%; Score 48; DB 1; Length 861;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||  
Db 674 LMNMFNI 680

RESULT 8  
env polypeptide precursor - human immunodeficiency virus type 1 (isolate SC)

N:Alternate names: coat polypeptide

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 16-Feb-1997

C:Accession: B28922

R:Gurgo, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Sta

Virolgy 164, 531-536, 1988

A:Title: Envelope sequences of two new United States HIV-1 isolates.

A:Reference number: A28922; MUID:88219542; PMID:3369091

A:Accession: B28922

A:Molecule type: DNA

A:Residues: 1-861 <GUR>

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polypeptide

C:Keywords: capsid protein; coat protein; glycoprotein; polypeptide; transmembrane prote

F:1.29/Domain: signal sequence #status predicted <SIG>

F:30-861/Product: env polypeptide #status predicted <APP>

F:87,129,135,140,143,159,163,167,198,234,241,262,276,295,301,302,333,340,356,362,386,396

Query Match 100.0%; Score 48; DB 1; Length 861;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||  
Db 674 LMNMFNI 680

RESULT 9  
env polypeptide precursor - human immunodeficiency virus type 1 (strain KB-2) (fragment)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 07-Feb-1997

C:Accession: J00266

R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.

submitted to JIPID, October 1991

A:Reference number: J00954

A:Accession: J00266

A:Molecule type: DNA

A:Residues: 1-136 <SHI>

C:Superfamily: type E retrovirus env polypeptide

C:Keywords: polypeptide

Query Match 89.6%; Score 43; DB 2; Length 136;  
Best Local Similarity 85.7%; Pred. No. 4.1;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||  
Db 58 LMNMFNI 64

RESULT 10  
env polypeptide precursor - human immunodeficiency virus type 1 (strain KB-19p41)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 07-Feb-1997

C:Accession: J00954

R:Shimizu, H.; Hasebe, F.; Tsuchie, H.; Morikawa, S.; Ushijima, H.; Kitamura, T.

submitted to JIPID, October 1991

A:Reference number: J00954

A:Accession: J00954

A:Molecule type: DNA

A:Residues: 1-136 <SHI>

C:Superfamily: type E retrovirus env polypeptide

C:Keywords: polypeptide

Query Match 89.6%; Score 43; DB 2; Length 136;  
Best Local Similarity 85.7%; Pred. No. 4.1;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||  
Db 58 LMNMFNI 64

RESULT 11  
env polypeptide precursor - human immunodeficiency virus type 1 (patient 4L)

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S70420; S22006

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer

A:Reference number: S70417; MUID:92144209; PMID:1736940

A:Accession: S70420

A:Molecule type: DNA

A:Status: preliminary

A:Residues: 1-357 <STR>

A:Cross-references: EMBL:X61354; NID:960190; PIDN:CAA43620.1; PID:960191

A:Experimental source: patient L

A:Note: submitted to the EMBL Data Library, July 1991

C:Superfamily: type E retrovirus env polypeptide

Query Match 89.6%; Score 43; DB 2; Length 357;  
Best Local Similarity 85.7%; Pred. No. 11;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||  
Db 170 LMNMFNI 176

RESULT 12  
env polypeptide precursor - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

A:Variety: isolate 4B

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000

C:Accession: S22004; S70419

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

submitted to the EMBL Data Library, July 1991

A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as dete

A:Reference number: S21950

A:Accession: S22004

A:Molecule type: DNA

A:Residues: 1-357 <STE1>

A:Cross-references: EMBL:X61353; NID:960188; PIDN:CAA43618.1; PID:960189

R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.

AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer

A:Reference number: S70417; MUID:92144209; PMID:1736940

A:Accession: S70419

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-292, 'X', 294-357 <STE2>

A:Cross-references: EMBL:X61353; NID:960188

C:Superfamily: type E retrovirus env polypeptide

Db 665 LMNMFNI 671

RESULT 3

S54384

envelope polyprotein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 26-Aug-1999

C:Accession: S54384

R:Theodore, T.; Buckler-White, A.J.  
submitted to the EMBL Data Library, July 1989

A:Reference number: S54377

A:Accession: S54384

A:Status: Preliminary

A:Molecule type: genomic RNA

A:Residues: 1-853 <THE>

A:Cross-references: EMBL:M2639; NID:g329377; PIDN:AAA5370.1; PID:g329385

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: polyprotein

Query Match 100.0%; Score 48; DB 2; Length 853;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||||

Db 666 LMNMFNI 672

RESULT 4

S13288

env protein - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997

C:Accession: S13288

R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, J.A.  
Nature 348, 69-73, 1990

A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120

A:Reference number: S13288; MUID:91043044; PMID:2172833

A:Accession: S13288

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-854 <OBR>

C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 48; DB 2; Length 854;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||||

Db 667 LMNMFNI 673

RESULT 5

VCLJZR

env polyprotein precursor - human immunodeficiency virus Zr-6

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus Zr-6

C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999

C:Accession: D26192

R:Striyvasan, A.; Anand, R.; York, D.; Ranganathan, P.; Feorjino, P.; Schochetman, G.; Cu  
Gene 52, 71-82, 1987

A:Title: Molecular characterization of human immunodeficiency virus from Zaire: nucleot  
A:Reference number: A26192; MUID:87248097; PMID:3036660

A:Accession: D26192

A:Molecule type: DNA

A:Residues: 1-855 <SRI>

A:Cross-references: GB:K03458; GB:M16322; NID:g329398; PIDN:AAA5380.1; PID:g329403

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polypro

F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-855/Product: env polyprotein #status predicted <MAT>  
F:20-500/Product: exterior membrane glycoprotein #status predicted <EXT>  
F:501-855/Product: transmembrane glycoprotein #status predicted <TM>  
F:87,129,140,145,154,158,186,189,199,236,243,264,278,281,297,333,340,355,386,392,398,

Query Match 100.0%; Score 48; DB 1; Length 855;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||||

Db 668 LMNMFNI 674

RESULT 6

VCLJH3

env polyprotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C:Accession: A03973

R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; D  
nberger, J.A.; Papas, T.S.; Graybe, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.  
Nature 313, 277-284, 1985

A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.

A:Reference number: A93353; MUID:85111123; PMID:2578615

A:Accession: A03973

A:Molecule type: DNA

A:Residues: 1-856 <RAT>

A:Cross-references: GB:M16554; GB:K02008; GB:K02009; GB:K02010; NID:g326383; PIDN:AAA

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; poly

F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-511/Product: exterior membrane glycoprotein #status predicted <TM>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status pre

Query Match 100.0%; Score 48; DB 1; Length 856;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||||

Db 669 LMNMFNI 675

RESULT 7

VCLJIV

env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)

N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1

A:Note: host Homo sapiens (man)

C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999

C:Accession: A03975

R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.  
Cell 40, 9-17, 1985

A:Title: Nucleotide sequence of the AIDS virus, LAV.

A:Reference number: A90866; MUID:85099333; PMID:2981635

A:Accession: A03975

A:Molecule type: DNA

A:Residues: 1-861 <RAI>

A:Cross-references: GB:K02013; NID:g326417; PIDN:AAB59751.1; PID:g326424

C:Genetics:

A:Gene: env

C:Superfamily: type E retrovirus env polyprotein

C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; poly

F:1-30/Domain: signal sequence #status predicted <SIG>  
F:31-516/Product: exterior membrane glycoprotein #status predicted <EXT>

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:36 ; Search time 2.24138 seconds  
(without alignments)  
300.235 Million cell updates/sec

Title: US-09-877-606-7  
Perfect score: 48  
Sequence: 1 LMNMFNI 7

Scoring table: BL0SUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	100.0	851	2 S33985	env polyprotein -
2	48	100.0	852	1 VCLJBR	env polyprotein -
3	48	100.0	853	2 S54384	env polyprotein -
4	48	100.0	854	2 S13288	env polyprotein -
5	48	100.0	855	1 VCLJZR	env polyprotein -
6	48	100.0	856	1 VCLJH3	env polyprotein -
7	48	100.0	861	1 VCLJIV	env polyprotein -
8	48	100.0	861	1 VCLJSC	env polyprotein -
9	43	89.6	136	2 JU0266	env polyprotein -
10	43	89.6	136	2 JT0954	env polyprotein -
11	43	89.6	357	2 S22006	env polyprotein -
12	43	89.6	357	2 S22004	env polyprotein -
13	43	89.6	357	2 S21996	env polyprotein -
14	43	89.6	358	2 S21998	env polyprotein -
15	43	89.6	358	2 S22002	env polyprotein -
16	43	89.6	358	2 S22000	env polyprotein -
17	43	89.6	358	2 S20417	env polyprotein -
18	43	89.6	443	2 C41621	env polyprotein -
19	43	89.6	443	2 B41621	env polyprotein -
20	43	89.6	729	1 VCLJFX	env polyprotein -
21	43	89.6	843	1 H44001	env polyprotein -
22	43	89.6	846	1 VCLJND	env polyprotein -
23	43	89.6	847	2 T09448	env polyprotein -
24	43	89.6	847	2 S13289	env polyprotein -
25	43	89.6	854	1 VCLJST	env polyprotein -
26	43	89.6	855	1 VCLJAZ	env polyprotein -
27	43	89.6	856	1 VCLJ3W	env polyprotein -
28	43	89.6	856	1 A44963	env polyprotein -
29	43	89.6	859	1 VCLJMN	env polyprotein -

30	43	89.6	859	2 T01672	env polyprotein -
31	43	89.6	861	1 VCLJBR	env polyprotein -
32	42	87.5	856	1 VCLJVL	env polyprotein -
33	41	85.4	357	2 S21994	env polyprotein -
34	40	83.3	125	2 A11358	env polyprotein -
35	40	83.3	424	2 JC5921	env polyprotein -
36	39	81.2	193	2 T29706	env polyprotein -
37	39	81.2	241	2 S59377	env polyprotein -
38	39	81.2	322	2 T40718	env polyprotein -
39	39	81.2	371	2 A19698	env polyprotein -
40	39	81.2	376	2 A56480	env polyprotein -
41	39	81.2	394	2 A34417	env polyprotein -
42	39	81.2	439	2 T50688	env polyprotein -
43	39	81.2	574	1 A42950	env polyprotein -
44	38	79.2	323	2 AD0994	env polyprotein -
45	38	79.2	398	2 C82362	env polyprotein -

## ALIGNMENTS

RESULT 1  
S33985  
env polyprotein - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999  
C:Accession: S33985  
R:Carlini, F.  
submitted to the EMBL Data Library, November 1991  
A:Reference number: S33979  
A:Accession: S33985  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-851 <CAR>  
A:Cross-references: EMBL:Z11530; NID:960192; PIDN:CAA77628.1; PID:960199  
C:Superfamily: type E retrovirus env polyprotein

Query Match 100.0%; Score 48; DB 2; Length 851;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
Db 664 LMNMFNI 670

## RESULT 2

VCLJBR  
env polyprotein - human immunodeficiency virus type 1 (isolate BR)  
N:Alternate names: coat protein gp120; coat protein gp41  
C:Contigs: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1989 #sequence\_revision 31-Dec-1989 #text\_change 16-Feb-1997  
C:Accession: A31667  
R:Amad, R.; Thayer, R.; Stinivasan, A.; Nayyar, S.; Gardner, M.; Luciw, P.; Dandekar  
Virology 168, 79-89, 1989  
A:Title: Biological and molecular characterization of human immunodeficiency virus (H  
A:Reference number: A94389; MIMD:89085613; PMID:2789516  
A:Accession: A31667  
A:Molecule type: DNA  
A:Residues: 1-852 <ANA>  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: capsid protein; coat protein; polyprotein; transmembrane protein  
F:1-516/Product: coat protein gp120 #status predicted <CP1>  
F:517-852/Product: coat protein gp41 #status predicted <CP2>

Query Match 100.0%; Score 48; DB 1; Length 852;  
Best Local Similarity 100.0%; Pred. No. 4.4;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 LMNMFNI 7  
|||||||



Thu Mar 27 11:38:00 2003

us-09-877-606-7.rnpb

Page 6

Db 7585 UUGUGAAUUGG00UACCA0A 7605

Search completed: March 25, 2003, 20:14:59  
Job time : 11.8966 secs

```
; Sequence 52, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 52
; LENGTH: 3839
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Tat
US-10-003-035-52

Alignment Scores:
Pred. No.: 123          Length: 3839
Score: 48.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9          Indels: 0
                          Gaps: 0

US-09-877-606-7 (1-7) x US-10-003-035-52 (1-3839)
OY 1 LeuTripsanTrpPheasnlle 7
Db 3570 TTGTGGAATTGTTTACATA 3590

RESULT 13
; Sequence 54, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 4040
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Tat/Rev
US-10-003-035-54

Alignment Scores:
Pred. No.: 131          Length: 4040
Score: 48.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 9          Indels: 0
                          Gaps: 0

US-09-877-606-7 (1-7) x US-10-003-035-54 (1-4040)
OY 1 LeuTripsanTrpPheasnlle 7
Db 3570 TTGTGGAATTGTTTACATA 3590
```

```
RESULT 14
; Sequence 3, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Human Immunodeficiency Virus
; FEATURE:
; NAME/KEY: source
; LOCATION: (1)...(8933)
; OTHER INFORMATION: Sequence of transcripts produced from the BH10
US-09-943-286-3

Alignment Scores:
Pred. No.: 326          Length: 8933
Score: 48.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10         Indels: 0
                          Gaps: 0

US-09-877-606-7 (1-7) x US-09-943-286-3 (1-8933)
OY 1 LeuTripsanTrpPheasnlle 7
Db 7585 UUGUGGAUUGGUUUAACAA 7605

RESULT 15
; Sequence 4, Application US/09943286
; Patent No. US2002010668A1
; GENERAL INFORMATION:
; APPLICANT: Nunomura, Kiyotada
; TITLE OF INVENTION: POLYNUCLEOTIDE AMPLIFICATION METHOD
; FILE REFERENCE: GP104-02.UT
; CURRENT APPLICATION NUMBER: US/09/943,286
; CURRENT FILING DATE: 2001-08-30
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 8933
; TYPE: RNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Sequence of the IAC-Asrc pseudo target
; NAME/KEY: mutation
; LOCATION: (4135)...(4155)
; OTHER INFORMATION: Mutated positions: 4135, 4140-1, 4145, 4150,
; OTHER INFORMATION: 4152-3, 4155
US-09-943-286-4

Alignment Scores:
Pred. No.: 326          Length: 8933
Score: 48.00           Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10         Indels: 0
                          Gaps: 0

US-09-877-606-7 (1-7) x US-09-943-286-4 (1-8933)
OY 1 LeuTripsanTrpPheasnlle 7
```

```
; CURRENT APPLICATION NUMBER: US/10/003.035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 2941
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env from HIV strain pNL4-3
US-10-003-035-19

Alignment Scores:
Pred. No.: 90.5 Length: 2941
Score: 48.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-877-606-7 (1-7) x US-10-003-035-19 (1-2941)

QY 1 LeuTrpAsnTrpPheAsnIle 7
Db 2471 TTGTGGAATTGTTTAAACATA 2491

RESULT 9
US-10-003-035-22
; Sequence 22, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003.035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 2950
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Nef from strain BH10
US-10-003-035-22

Alignment Scores:
Pred. No.: 90.8 Length: 2950
Score: 48.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-877-606-7 (1-7) x US-10-003-035-22 (1-2950)

QY 1 LeuTrpAsnTrpPheAsnIle 7
Db 2011 TTGTGGAATTGTTTAAACATA 2031

RESULT 10
US-10-003-035-14
; Sequence 14, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003.035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 3157
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Env/Tat/Rev
US-10-003-035-14

Alignment Scores:
Pred. No.: 98.2 Length: 3157
Score: 48.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-877-606-7 (1-7) x US-10-003-035-14 (1-3157)

QY 1 LeuTrpAsnTrpPheAsnIle 7
Db 2484 TTGTGGAATTGTTTAAACATA 2504

RESULT 11
US-10-003-035-21
; Sequence 21, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003.035
; CURRENT FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 3417
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Tat/Rev/Nef from strain BH10
US-10-003-035-21

Alignment Scores:
Pred. No.: 108 Length: 3417
Score: 48.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-09-877-606-7 (1-7) x US-10-003-035-21 (1-3417)

QY 1 LeuTrpAsnTrpPheAsnIle 7
Db 2478 TTGTGGAATTGTTTAAACATA 2498

RESULT 12
US-10-003-035-52
```

```

; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 32
; LENGTH: 2562
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Env with multi-clade V3 loops
US-10-003-035-32

```

```

Alignment Scores:
Pred. No.: 77.1      Length: 2562
Score: 48.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9                Gaps: 0

```

US-09-877-606-7 (1-7) x US-10-003-035-32 (1-2562)

```

OY 1 LeutPrasnrPpHeasnlle 7
Db 2452 TTGTGGAATTGCTTACATA 2472

```

```

RESULT 5
US-10-003-035-24
; Sequence 24, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 24
; LENGTH: 2583
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env
US-10-003-035-24

```

```

Alignment Scores:
Pred. No.: 77.9      Length: 2583
Score: 48.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9                Gaps: 0

```

US-09-877-606-7 (1-7) x US-10-003-035-24 (1-2583)

```

OY 1 LeutPrasnrPpHeasnlle 7
Db 2017 TTGTGGAATTGCTTACATA 2037

```

```

RESULT 6
US-10-003-035-20
; Sequence 20, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599

```

```

; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 20
; LENGTH: 2746
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Tat/Rev from pML4-3
US-10-003-035-20

```

```

Alignment Scores:
Pred. No.: 83.6      Length: 2746
Score: 48.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9                Gaps: 0

```

US-09-877-606-7 (1-7) x US-10-003-035-20 (1-2746)

```

OY 1 LeutPrasnrPpHeasnlle 7
Db 2276 TTGTGGAATTGCTTACATA 2296

```

```

RESULT 7
US-10-003-035-23
; Sequence 23, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 23
; LENGTH: 2747
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env/Tat from strain BH10
US-10-003-035-23

```

```

Alignment Scores:
Pred. No.: 83.6      Length: 2747
Score: 48.00         Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9                Gaps: 0

```

US-09-877-606-7 (1-7) x US-10-003-035-23 (1-2747)

```

OY 1 LeutPrasnrPpHeasnlle 7
Db 2478 TTGTGGAATTGCTTACATA 2498

```

```

RESULT 8
US-10-003-035-19
; Sequence 19, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712

```

```
; FILING DATE: 17-AUG-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/987,867
; FILING DATE: 09-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/389,459
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Lauro, Peter C.
; REGISTRATION NUMBER: 32,360
; REFERENCE/DOCKET NUMBER: UAI-004CPDVZCN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1568 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 7..1565
; US-09-756-551A-7

Alignment Scores:
Pred. No.: 43.8      Length: 1568
Score: 48.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 10      Gaps: 0

US-09-877-606-7 (1-7) x US-09-756-551A-7 (1-1568)
Oy 1 LeuTrpAsnTrpPheAsnIle 7
    |||||||
Db 1399 TTGTGGAATTGCTTTAACATA 1419

RESULT 2
US-10-003-035-16
; Sequence 16, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 2280
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Modified Env
US-10-003-035-16

Alignment Scores:
Pred. No.: 67.4      Length: 2280
Score: 48.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
```

```
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 9      Gaps: 0

US-09-877-606-7 (1-7) x US-10-003-035-16 (1-2280)
Oy 1 LeuTrpAsnTrpPheAsnIle 7
    |||||||
Db 2011 TTGTGGAATTGCTTTAACATA 2031

RESULT 3
US-09-991-258-18
; Sequence 18, Application US/09991258
; Patent No. US20020141975A1
; GENERAL INFORMATION:
; APPLICANT: Olmsted, Robert
; APPLICANT: Keith, Paula
; APPLICANT: Dryga, Sergey
; APPLICANT: Caley, Ian
; APPLICANT: Maughan, Maureen
; APPLICANT: Johnston, Robert
; APPLICANT: Davis, Nancy
; APPLICANT: Swanstrom, Ronald
; TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR U
; FILE REFERENCE: 01113.0001U3
; CURRENT APPLICATION NUMBER: US/09/991,258
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 09/902,537
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/216,995
; PRIOR FILING DATE: 2000-07-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18
; LENGTH: 2532
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; NO. US20020141975A1e -
; NAME/KEY: CDS
; LOCATION: (1)...(2532)
US-09-991-258-18

Alignment Scores:
Pred. No.: 76.1      Length: 2532
Score: 48.00      Matches: 7
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match: 100.00%      Indels: 0
DB: 10      Gaps: 0

US-09-877-606-7 (1-7) x US-09-991-258-18 (1-2532)
Oy 1 LeuTrpAsnTrpPheAsnIle 7
    |||||||
Db 1948 TTGTGGAATTGCTTTAACATA 1968

RESULT 4
US-10-003-035-32
; Sequence 32, Application US/10003035
; Patent No. US20020155127A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Danher
; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS
; FILE REFERENCE: 22488-712
; CURRENT APPLICATION NUMBER: US/10/003,035
; PRIOR FILING DATE: 2001-11-01
; PRIOR APPLICATION NUMBER: 09/585,599
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US01/18238
; PRIOR FILING DATE: 2001-06-04
```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 15:49:42 ; Search time 9.89655 Seconds  
(without alignments)  
548.507 Million cell updates/sec

Title: US-09-877-606-7  
Perfect score: 48  
Sequence: 1 LMNMENI 7

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 538826 seqs, 387737923 residues

Total number of hits satisfying chosen parameters: 1077652

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame\_plus\_p2n.model -DEV=x1h  
-O/cgn2.1/USPTO.spool/US09877606/runat\_18032003\_134802\_17783/app\_query.fasta.1.1059  
-DB=Published.Applications\_NA\_QPMT=fastap -SUFFIX=trnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=BLOSUM62  
-TRANS=human40.csl -LIST=45 -DOCCALIGN=200 -THR\_SCORE=pcr -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NCRM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09877606.gcg.1.1.115/runat\_18032003\_134802\_17783  
-NCPU=6 -ICPU=3 -NO\_XLPRY -NO\_MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -DELOP=6 -DELEXT=7

Database: Published.Applications\_NA:\*

1: /cgn2.6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2.6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2.6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2.6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2.6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2.6/ptodata/2/pubpna/PCUS\_PUBCOMB.seq:\*  
7: /cgn2.6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2.6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2.6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
10: /cgn2.6/ptodata/2/pubpna/US09\_PUBCOMB.seq:\*  
11: /cgn2.6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
12: /cgn2.6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
13: /cgn2.6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
14: /cgn2.6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	1568	10	US-09-756-551A-7
2	48	100.0	2280	9	US-10-003-035-16
3	48	100.0	2352	10	US-09-991-258-18
4	48	100.0	2562	9	US-10-003-035-32

5	48	100.0	2583	9	US-10-003-035-24	Sequence 24, Appl
6	48	100.0	2746	9	US-10-003-035-20	Sequence 20, Appl
7	48	100.0	2747	9	US-10-003-035-23	Sequence 23, Appl
8	48	100.0	2941	9	US-10-003-035-19	Sequence 19, Appl
9	48	100.0	2950	9	US-10-003-035-22	Sequence 22, Appl
10	48	100.0	3157	9	US-10-003-035-14	Sequence 14, Appl
11	48	100.0	3417	9	US-10-003-035-21	Sequence 21, Appl
12	48	100.0	3839	9	US-10-003-035-52	Sequence 52, Appl
13	48	100.0	4040	9	US-10-003-035-54	Sequence 54, Appl
14	48	100.0	8933	10	US-09-943-286-3	Sequence 3, Appl
15	48	100.0	8933	10	US-09-943-286-4	Sequence 4, Appl
16	48	100.0	8933	10	US-09-943-286-9	Sequence 9, Appl
17	48	100.0	9719	10	US-09-737-190A-1	Sequence 1, Appl
18	48	100.0	9918	10	US-09-798-675-5	Sequence 5, Appl
19	48	100.0	13584	10	US-09-991-258-17	Sequence 17, Appl
20	44	91.7	388	10	US-09-878-574-1625	Sequence 1625, Ap
21	44	91.7	389	10	US-09-878-574-1625	Sequence 1625, Ap
22	44	91.7	400	10	US-09-878-574-1648	Sequence 1648, Ap
23	44	91.7	451	10	US-09-960-352-1342	Sequence 1342, Ap
24	44	91.7	3063	10	US-09-815-242-4416	Sequence 4416, Ap
25	44	91.7	3066	10	US-09-815-242-8446	Sequence 8446, Ap
26	44	91.7	32463	9	US-09-996-956-5	Sequence 5, Appl
27	43	89.6	1057	9	US-09-886-156-44	Sequence 44, Appl
28	43	89.6	1057	9	US-09-886-156-45	Sequence 45, Appl
29	43	89.6	1057	9	US-09-886-150-44	Sequence 44, Appl
30	43	89.6	1057	9	US-09-886-150-45	Sequence 45, Appl
31	43	89.6	1057	9	US-09-886-149-44	Sequence 44, Appl
32	43	89.6	1057	9	US-09-886-149-45	Sequence 45, Appl
33	43	89.6	1057	9	US-09-886-159-44	Sequence 44, Appl
34	43	89.6	1057	9	US-09-886-159-45	Sequence 45, Appl
35	43	89.6	1860	10	US-09-891-609-3	Sequence 3, Appl
36	43	89.6	1941	10	US-09-891-609-1	Sequence 1, Appl
37	43	89.6	2298	10	US-09-476-242-22	Sequence 22, Appl
38	43	89.6	2298	10	US-09-476-242-23	Sequence 23, Appl
39	43	89.6	2298	10	US-09-476-242-24	Sequence 24, Appl
40	43	89.6	2310	10	US-09-476-242-3	Sequence 3, Appl
41	43	89.6	2310	10	US-09-476-242-21	Sequence 21, Appl
42	43	89.6	2316	10	US-09-476-242-4	Sequence 4, Appl
43	43	89.6	2316	10	US-09-476-242-8	Sequence 8, Appl
44	43	89.6	2322	10	US-09-476-242-5	Sequence 5, Appl
45	43	89.6	2322	10	US-09-476-242-18	Sequence 18, Appl

#### ALIGNMENTS

RESULT 1  
US-09-756-551A-7  
Sequence 7, Application US/09756551A  
Patent No. US20020051768A1  
GENERAL INFORMATION:  
APPLICANT: C. Morrow et al.  
TITLE OF INVENTION: ENCAPSATED RECOMBINANT VIRAL  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
TITLE OF INVENTION: USING SAME  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 STATE STREET  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/756,551A  
FILING DATE: 08-JAN-2001  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/376,184

**THIS PAGE BLANK (USPTO)**

LOCATION: 12..2600  
US-08-472-240A-9  
Alignment Scores:  
Pred. No.: 56.3  
Score: 48.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 4  
US-09-877-606-7 (1-7) x US-08-472-240A-9 (1-2644)  
Qy 1 Leutrpasnrppheasnille 7  
Db 2031 TTGTGGAATTGCTTACATA 2051  
RESULT 14  
US-08-147-890-1  
Sequence 1, Application US/08147890  
Patent No. 5714316  
GENERAL INFORMATION:  
APPLICANT: Weiner, David  
APPLICANT: Williams, William  
APPLICANT: Levy, David N.  
TITLE OF INVENTION: Chimeric Envelope Proteins for Viral  
TITLE OF INVENTION: Targeting  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Howson and Howson  
STREET: Spring House Corporate Cntr, P.O. Box 457  
CITY: Spring House  
STATE: Pennsylvania  
COUNTRY: USA  
ZIP: 19477  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/147,890  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/718,537  
FILING DATE: 21-JUN-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bak, Mary E.  
REGISTRATION NUMBER: 31,215  
REFERENCE/DOCKET NUMBER: WST15AUSA  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 215-540-9200  
TELEFAX: 215-540-5818  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2694 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: DNA (genomic)  
US-08-147-890-1  
Alignment Scores:  
Pred. No.: 57.4  
Score: 48.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 1  
US-09-877-606-7 (1-7) x US-08-147-890-1 (1-2694)

Qy 1 Leutrpasnrppheasnille 7  
Db 2113 TTGTGGAATTGCTTACATA 2133  
RESULT 15  
US-08-728-122-1  
Sequence 1, Application US/08728122  
Patent No. 6074646  
GENERAL INFORMATION:  
APPLICANT: Cloyd, Miles W.  
APPLICANT: Ramsey, Keith W.  
TITLE OF INVENTION: A New EIA Test Using No. 6074646-Denatured  
TITLE OF INVENTION: HIV Antigen for Early Detection of  
TITLE OF INVENTION: HIV Infection  
NUMBER OF SEQUENCES: 1  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Akin, Gump, Strauss, Hauer & Feld, L.L.P.  
STREET: 816 Congress Avenue, Suite 1900  
CITY: Austin  
STATE: Texas  
COUNTRY: USA  
ZIP: 78701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/728,122  
FILING DATE: 09-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Mayfield, Denise L.  
REGISTRATION NUMBER: 33,732  
REFERENCE/DOCKET NUMBER: 43424.0003  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 499-6200  
TELEFAX: (512) 499-6290  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2730 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-728-122-1  
Alignment Scores:  
Pred. No.: 58.2  
Score: 48.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 3  
US-09-877-606-7 (1-7) x US-08-728-122-1 (1-2730)  
Qy 1 Leutrpasnrppheasnille 7  
Db 2474 TTGTGGAATTGCTTACATA 2494  
Search completed: March 25, 2003, 15:52:20  
Job time : 6.37931 secs



US-07-956-483-18

Alignment Scores:

Pred. No.:	53.6	Length:	2531
Score:	48.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-877-606-7 (1-7) x US-07-956-483-18 (1-2531)

QY 1 Leutrpasnrppheasnlle 7  
|||||

Db 2031 TTGTGATGTTTACATA 2051

RESULT 12

US-07-916-098A-1

Sequence 1, Application US/07916098A

Patent No. 5871732

GENERAL INFORMATION:

APPLICANT: BURKLY, LINDA C.

APPLICANT: THOMAS, PATRICIA L.

APPLICANT: ROSA, MARGARET D.

APPLICANT: ROSA, JOSEPH J.

TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN

TITLE OF INVENTION: PROPHYLAXIS AND TREATMENT OF AIDS, ARC AND HIV INFECTION

NUMBER OF SEQUENCES: 61

CORRESPONDENCE ADDRESS:

ADDRESSEE: ALLEGRETTI & MITCOFF, LTD.

STREET: 10 SOUTH WACKER DRIVE

CITY: CHICAGO

STATE: ILLINOIS

COUNTRY: U.S.A.

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WORD PERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/07/916,098A

FILING DATE: JULY 24, 1992

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/08843

FILING DATE: NO. 5871732ember 27, 1991

CLASSIFICATION: 424

APPLICATION NUMBER: 07/618,542

FILING DATE: NO. 5871732ember 27, 1990

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: JOHN J. MC DONNELL

REGISTRATION NUMBER: 26,949

REFERENCE/DOCKET NUMBER: 92,310-G

TELECOMMUNICATION INFORMATION:

TELEPHONE: (312) 715-1000

TELEFAX: (312) 715-1234

TELEX: 910/221-5317

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2571 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

HYPOTHETICAL: NO

ANTI-SENSE: NO

FEATURE:

NAME/KEY: sig\_peptide

LOCATION: 1..87

FEATURE:

NAME/KEY: mat\_peptide

LOCATION: 88..2568

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1533..1534

OTHER INFORMATION: /note= "gp120/gp41 cleavage site"

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2568

FEATURE:

NAME/KEY: misc\_feature

LOCATION: 1

OTHER INFORMATION: /note= "pre-HIV gp160"

US-07-916-098A-1

Alignment Scores:

Pred. No.:	54.6	Length:	2571
Score:	48.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	2	Gaps:	0

US-09-877-606-7 (1-7) x US-07-916-098A-1 (1-2571)

QY 1 Leutrpasnrppheasnlle 7  
|||||

Db 2005 TTGTGATGTTTACATA 2025

RESULT 13

US-08-472-240A-9

Sequence 9, Application US/08472240A

Patent No. 6284248

GENERAL INFORMATION:

APPLICANT: KIENY, Marie-Paule

TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED

TITLE OF INVENTION: gp160 VARIANT

NUMBER OF SEQUENCES: 29

CORRESPONDENCE ADDRESS:

ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS

STREET: P.O. Box 1404

CITY: Alexandria

STATE: Virginia

COUNTRY: United States

ZIP: 22313-1404

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/472,240A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/956,483

FILING DATE: 31-DEC-1992

ATTORNEY/AGENT INFORMATION:

NAME: Teskin, Robin L.

REGISTRATION NUMBER: 35,030

REFERENCE/DOCKET NUMBER: 017753-055

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 836-6620

TELEFAX: (703) 836-2021

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 2644 base pairs

TYPE: nucleic acid

STRANDEDNESS: not relevant

TOPOLOGY: not relevant

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

APPLICANT: Nicholas J. Deacon  
APPLICANT: Dale A. McPhee  
APPLICANT: David Cooper  
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
NUMBER OF SEQUENCES: 841  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCULLY, SCOTT, MURPHY & PRESSER  
STREET: 400 GARDEN CITY PLAZA  
CITY: GARDEN CITY  
STATE: NEW YORK  
COUNTRY: U.S.A.  
ZIP: 11530-0299  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/488,551B  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PM3864 (AU)  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: PM4002 (AU)  
FILING DATE: 21-FEB-1994  
APPLICATION NUMBER: PM0284 (AU)  
FILING DATE: 23-DEC-1994  
APPLICATION NUMBER: US 08/388,353  
FILING DATE: 14-FEB-1995  
APPLICATION NUMBER: PM3021/95  
FILING DATE: 17-MAY-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: FRANK S. DIGILIO  
REFERENCE/DOCKET NUMBER: 96062  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
INFORMATION FOR SEQ ID NO: 651:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1596 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
US-08-488-551B-651

Alignment Scores:  
Pred. No.: 32.5 Length: 1596  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 3 Gaps: 0

US-09-877-606-7 (1-7) x US-08-488-551B-651 (1-1596)  
Oy 1 LeuTpsantrpPheasnlle 7  
Db 148 TTGTGGAATTGTTTAACATA 168

RESULT 10  
US-09-337-387-4  
Sequence 4, Application US/09337387  
Patent No. 6420545  
GENERAL INFORMATION:  
APPLICANT: HOXIE, James A.  
APPLICANT: LABRANCHE, Celia C.  
APPLICANT: DOMS, Robert W.  
APPLICANT: HOFFMAN, Trevor L.  
TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND  
FILE REFERENCE: Hoxie 9596-104U1 (0282)  
CURRENT APPLICATION NUMBER: US/09/337,387

CURRENT FILING DATE: 1999-06-22  
PRIOR APPLICATION NUMBER: US 09/317,556  
PRIOR FILING DATE: 1999-05-24  
NUMBER OF SEQ ID NOS: 12  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 2184  
TYPE: DNA  
ORGANISM: Human immunodeficiency virus type 1  
US-09-337-387-4

Alignment Scores:  
Pred. No.: 45.7 Length: 2184  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 4 Gaps: 0

US-09-877-606-7 (1-7) x US-09-337-387-4 (1-2184)  
Oy 1 LeuTpsantrpPheasnlle 7  
Db 1990 TTGTGGAATTGTTTAACATA 2010

RESULT 11  
US-07-956-483-18  
Sequence 18, Application US/07956483  
Patent No. 6261799  
GENERAL INFORMATION:  
APPLICANT: KIENY, Marie-Paule  
TITLE OF INVENTION: NOVEL HYBRID, SOLUBLE AND UNCLEAVED  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Burns, Doane, Swecker & Mathis  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/956,483  
FILING DATE: 31-DEC-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO 92/19742  
FILING DATE: 12-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: FR 91 05392  
FILING DATE: 02-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Crane-Feurty, Sharon E  
REGISTRATION NUMBER: 36,113  
REFERENCE/DOCKET NUMBER: 017753-005  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2531 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 12..2531

MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7..1565  
US-08-389-459A-7

Alignment Scores:  
Pred. No.: 31.9  
Score: 48.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 1  
Length: 1568  
Matches: 7  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-877-606-7 (1-7) x US-08-389-459A-7 (1-1568)

OY 1 Leutrpasntppheasnlle 7  
Db 1399 TTGTGAATGTTTACATA 1419

RESULT 7  
US-08-987-867A-7  
Sequence 7, Application US/08987867A  
Patent No. 6063384  
GENERAL INFORMATION:  
APPLICANT: C. Morrow et al.  
TITLE OF INVENTION: ENCAPSATED RECOMBINANT VIRAL  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 STATE STREET  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/987,867A  
FILING DATE: 09-DEC-1997  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/087,009  
FILING DATE: 01-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Myers, Louis  
REGISTRATION NUMBER: 35,965  
REFERENCE/DOCKET NUMBER: UAP-004CPDV  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1568 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7..1565  
US-08-987-867A-7

Alignment Scores:  
Pred. No.: 31.9  
Score: 48.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Length: 1568  
Matches: 7  
Conservative: 0  
Mismatch: 0  
Indels: 0

Query Match: 100.00%  
DB: 3  
Indels: 0  
Gaps: 0

US-09-877-606-7 (1-7) x US-08-987-867A-7 (1-1568)

OY 1 Leutrpasntppheasnlle 7  
Db 1399 TTGTGAATGTTTACATA 1419

RESULT 8  
US-08-388-353-651  
Sequence 651, Application US/08388353  
Patent No. 6010895  
GENERAL INFORMATION:  
APPLICANT: Deacon, Nicholas J.  
APPLICANT: Learmont, Jennifer C.  
APPLICANT: McPhee, Dale A.  
APPLICANT: Crowe, Suzanne  
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
NUMBER OF SEQUENCES: 800  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States  
ZIP: 11530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/388,353  
FILING DATE: 14-FEB-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Digiglio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9606  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR  
INFORMATION FOR SEQ ID NO: 651:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1596 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-388-353-651

Alignment Scores:  
Pred. No.: 32.5  
Score: 48.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 3  
Length: 1596  
Matches: 7  
Conservative: 0  
Mismatch: 0  
Indels: 0  
Gaps: 0

US-09-877-606-7 (1-7) x US-08-388-353-651 (1-1596)

OY 1 Leutrpasntppheasnlle 7  
Db 148 TTGTGAATGTTTACATA 168

RESULT 9  
US-08-488-551B-651  
Sequence 651, Application US/08488551B  
Patent No. 6015661  
GENERAL INFORMATION:

SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/589,446  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/087,009  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Geary III, William C.  
REGISTRATION NUMBER: 31,359  
REFERENCE/DOCKET NUMBER: UAG-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1568 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7..1565  
US-08-589-446-7

Alignment Scores:  
Pred. No. 31.9 Length: 1568  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-877-606-7 (1-7) x US-08-589-446-7 (1-1568)  
QY 1 LeuTriPaSnTrpPheAsnIle 7  
Db 1399 TTGTGGAATTGTTTACATA 1419

RESULT 5  
US-08-444-882-7  
Sequence 7, Application US/08444882  
Patent No. 5622705  
GENERAL INFORMATION:  
APPLICANT: Morrow, Casey D.  
TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC  
TITLE OF INVENTION: ACID AND METHODS OF MAKING AND  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/444,882  
FILING DATE: 19-MAY-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/087,009  
FILING DATE: 01-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Geary III, William C.  
REGISTRATION NUMBER: 31,359

REFERENCE/DOCKET NUMBER: UAG-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1568 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 7..1565  
US-08-444-882-7

Alignment Scores:  
Pred. No. 31.9 Length: 1568  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 1 Gaps: 0

US-09-877-606-7 (1-7) x US-08-444-882-7 (1-1568)  
QY 1 LeuTriPaSnTrpPheAsnIle 7  
Db 1399 TTGTGGAATTGTTTACATA 1419

RESULT 6  
US-08-389-459A-7  
Sequence 7, Application US/08389459A  
Patent No. 5817512  
GENERAL INFORMATION:  
APPLICANT: Morrow, Casey D. and Porter, Donna, C.  
TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT POLIOVIRUS  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/389,459A  
FILING DATE: 15-FEB-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/087,009  
FILING DATE: 01-JUL-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Silverl, Jean M.  
REGISTRATION NUMBER: 39,030  
REFERENCE/DOCKET NUMBER: UAG-004CP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1568 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear

QY 1 LeuTpaSntRpheaSnile 7  
|||||  
Db 523 TTGTGAATGTGTTAACATA 543

## RESULT 2

US-08-467-933-7  
; Sequence 7, Application US/08467933  
; Patent No. 5736391  
; GENERAL INFORMATION:  
; APPLICANT: Essex, Myron E.  
; APPLICANT: Yu, Xiaofang  
; APPLICANT: Lee, Tun-Hou  
; TITLE OF INVENTION: GP41 MUTANTS AND THEIR USE AS  
; TITLE OF INVENTION: HIV THERAPEUTICS  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: MS-DOS (Version 5.0)  
; SOFTWARE: Wordperfect (Version 5.1)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/467,933  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/979,975  
; FILING DATE: No. 5736391ember 23, 1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Freeman, John W.  
; REGISTRATION NUMBER: 29,066  
; REFERENCE/DOCKET NUMBER: 00379/019001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1260  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; US-08-467-933-7

Alignment Scores:  
Pred. No.: 25.1 Length: 1260  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-877-606-7 (1-7) x US-08-467-933-7 (1-1260)

QY 1 LeuTpaSntRpheaSnile 7  
|||||  
Db 599 TTGTGAATGTGTTAACATA 619

## RESULT 3

US-07-924-028A-1  
; Sequence 1, Application US/07924028A  
; Patent No. 5470573  
; GENERAL INFORMATION:  
; APPLICANT: Lubitz, Werner, Sczostak, Michael P.  
; TITLE OF INVENTION: CARRIER-BOUND RECOMBINANT PROTEINS, PROCESS  
; TITLE OF INVENTION: FOR THE PRODUCTION AND USE AS IMMUNOGENS AND VACCINES

NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Felfe & Lynch  
STREET: 805 Third Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10022

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 5.25 Inch, 360 kb storage  
COMPUTER: IBM PS/2  
OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/924,028A  
FILING DATE: 30-SEP-1992  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP91/00308  
FILING DATE: 02-FEB-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 40 05 874  
FILING DATE: 24-FEB-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hanson, No. 3470573man D.  
REGISTRATION NUMBER: 30,946  
REFERENCE/DOCKET NUMBER: HUBR 1027  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 688-9200  
TELEFAX: (212) 838-3884  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1451 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-07-924-028A-1

Alignment Scores:  
Pred. No.: 29.3 Length: 1451  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
Gaps: 0

US-09-877-606-7 (1-7) x US-07-924-028A-1 (1-1451)

QY 1 LeuTpaSntRpheaSnile 7  
|||||  
Db 631 TTGTGAATGTGTTAACATA 651

RESULT 4  
US-08-589-446-7  
; Sequence 7, Application US/08589446  
; Patent No. 5614413  
; GENERAL INFORMATION:  
; APPLICANT: Morrow, Casey D.  
; TITLE OF INVENTION: EMCAPSIDATED POLIOVIRUS NUCLEIC  
; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: LAHIVE & COCKFIELD  
; STREET: 60 STATE STREET, SUITE 510  
; CITY: BOSTON  
; STATE: MASSACHUSETTS  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - nucleic search, using frame\_plus.p2n model

Run on: March 25, 2003, 12:42:33 ; Search time 5.37931 seconds  
(without alignments)  
399.073 Million cell updates/sec

Title: US-09-877-606-7  
Perfect score: 48  
Sequence: 1 LMNMFNI 7

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-O=/cgn2.1/USFTO.spool/US09877606/runat\_18032003\_134801\_17715/app\_query.fasta.1.1059  
-DB=Issued\_Patents\_NA -QPM=fastap -SUFFIX=fnl -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human4.0.cdl  
-LIST=45 -DOCALLION=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USR=US09877606.cgn2.1.101 @runat\_18032003\_134801\_17715 -NCPU=6 -ICPU=3  
-NO\_XLPMY -NO\_MMP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEDOUT=120  
-WARN\_TIMEDOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- Issued\_Patents\_NA:\*
- 1: /cgn2.6/ptodata/1/ina/3A.COMB.seq:\*
  - 2: /cgn2.6/ptodata/1/ina/3B.COMB.seq:\*
  - 3: /cgn2.6/ptodata/1/ina/3A.COMB.seq:\*
  - 4: /cgn2.6/ptodata/1/ina/3B.COMB.seq:\*
  - 5: /cgn2.6/ptodata/1/ina/PCTUS.COMB.seq:\*
  - 6: /cgn2.6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	48	100.0	699	5	PCT-US95-13335-2
2	48	100.0	1260	1	US-08-467-933-7
3	48	100.0	1451	1	US-07-924-028A-1
4	48	100.0	1568	1	US-08-589-446-7
5	48	100.0	1568	1	US-08-444-882-7
6	48	100.0	1568	1	US-08-389-459A-7
7	48	100.0	1568	3	US-08-987-867A-7
8	48	100.0	1568	3	US-08-388-353-651
9	48	100.0	1596	3	US-08-488-551B-651
10	48	100.0	2184	4	US-09-337-387-4
11	48	100.0	2531	4	US-07-956-483-18
12	48	100.0	2571	2	US-07-916-098A-1

13	48	100.0	2644	4	US-08-472-240A-9	Sequence 9, Appl1
14	48	100.0	2694	1	US-08-147-890-1	Sequence 1, Appl1
15	48	100.0	2730	3	US-08-728-122-1	Sequence 1, Appl1
16	48	100.0	2945	6	5462872-1	Patent No. 5462872
17	48	100.0	3084	1	US-08-147-890-2	Sequence 2, Appl1
18	48	100.0	3563	3	US-08-463-210-6	Sequence 6, Appl1
19	48	100.0	7399	2	US-08-418-848A-9	Sequence 9, Appl1
20	48	100.0	8932	4	US-09-124-900-1	Sequence 1, Appl1
21	48	100.0	8933	3	US-08-463-210-4	Sequence 4, Appl1
22	48	100.0	8933	4	US-09-620-958A-3	Sequence 3, Appl1
23	48	100.0	8933	4	US-09-620-958A-4	Sequence 4, Appl1
24	48	100.0	8933	4	US-09-620-958A-9	Sequence 9, Appl1
25	48	100.0	9709	2	US-08-188-583-5	Sequence 5, Appl1
26	48	100.0	9709	3	US-08-388-353-1	Sequence 1, Appl1
27	48	100.0	9709	3	US-08-488-551B-1	Sequence 1, Appl1
28	48	100.0	9709	4	US-09-309-572-15	Sequence 15, Appl1
29	48	100.0	12494	4	US-08-935-112-13	Sequence 13, Appl1
30	48	100.0	12494	4	US-08-848-760B-33	Sequence 33, Appl1
31	48	100.0	15581	3	US-08-646-538-35	Sequence 35, Appl1
32	48	100.0	15581	4	US-09-503-222-35	Sequence 35, Appl1
33	44	91.7	51	1	US-08-323-192D-44	Sequence 44, Appl1
34	44	91.7	51	1	US-08-470-887A-43	Sequence 43, Appl1
35	44	91.7	51	2	US-08-252-508B-43	Sequence 43, Appl1
36	44	91.7	51	3	US-09-106-377-43	Sequence 43, Appl1
37	44	91.7	3498	3	US-08-293-728-1	Sequence 1, Appl1
38	44	91.7	3498	4	US-09-421-868-1	Sequence 1, Appl1
39	44	91.7	162450	4	US-09-345-882-1	Sequence 1, Appl1
40	43	89.6	1057	1	US-08-470-202-44	Sequence 44, Appl1
41	43	89.6	1057	1	US-08-471-770-44	Sequence 44, Appl1
42	43	89.6	1057	2	US-08-468-059-44	Sequence 44, Appl1
43	43	89.6	1057	4	US-09-109-916-44	Sequence 44, Appl1
44	43	89.6	1057	4	US-09-109-916-45	Sequence 45, Appl1
45	43	89.6	1208	3	US-08-388-353-615	Sequence 615, Appl1

#### ALIGNMENTS

RESULT 1  
PCT-US95-13335-2  
Sequence 2, Application PC/TUS9513335  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: A RECOMBINANT PROTEIN USEFUL IN THE  
DETECTION OF HIV, DNA SEQUENCES ENCODING THE PROTEIN, AND  
TITLE OF INVENTION: IMMUNOASSAYS USING THE PROTEIN.  
NUMBER OF SEQUENCES: 2  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/13335  
FILING DATE:  
CLASSIFICATION:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 699 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: unknown  
MOLECULE TYPE: cDNA  
PCT-US95-13335-2  
Alignment Scores:  
Pred. No.: 13.3  
Score: 48.00  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 100.00%  
DB: 5  
Gaps: 0  
Length: 699  
Matches: 7  
Conservative: 0  
Mismatch: 0  
Indels: 0  
US-09-877-606-7 (1-7) x PCT-US95-13335-2 (1-699)

XX 16-NOV-2000 (first entry)  
 DT  
 XX  
 DE sybX08- 41 gene derivative for use in HIV DNA vaccine.  
 XX  
 KW HIV; human immunodeficiency virus; vaccine; AIDS; snut;  
 KW silent nucleotide substitution; ds.  
 XX  
 OS Human immunodeficiency virus type 1.  
 OS Synthetic.  
 XX  
 PN WO200029561-A2.  
 PD  
 XX 25-MAY-2000.  
 PD  
 PF 27-MAR-2000; 2000WO-DK00144.  
 PF  
 PR 29-MAR-1999; 99DK-0000427.  
 PR 09-APR-1999; 99US-0128358.  
 XX  
 PA (STAT-) STATENS SERUM INST.  
 PA  
 PI Fomsgaard A;  
 DR WPI: 2000-387778/33.  
 DR P-PSDB: AAY99928.  
 XX  
 PT Producing nucleotide sequence construct with optimized codons for human  
 PT immunodeficiency virus (HIV) genetic vaccine involves obtaining a first  
 PT nucleotide sequence from a HIV patient, redesigning and assembling it  
 PT with snuts -  
 XX  
 PS Disclosure; Page 148-149; 150pp; English.  
 XX  
 CC The present invention relates to a nucleotide construct with optimised  
 CC codons for use as a human immunodeficiency virus (HIV) DNA vaccine. The  
 CC construct uses codons from highly expressed mammalian proteins to code  
 CC for each derivative of an early, primary HIV envelope gene. The first  
 CC stage in the production of the construct was the cloning of an HIV  
 CC envelope gene. A nucleotide sequence encoding this gene was then created  
 CC using codons from highly expressed mammalian genes. In the next  
 CC stage, snuts (AAA9060-A49079) were created by redesigning this  
 CC nucleotide construct so that restriction enzyme sites surrounded  
 CC functional regions of the sequence. The snuts were then assembled  
 CC to form thirteen pieces (AAA9080-A49092). The present sequence shows  
 CC a derivative of the envelope gene (AAA9093-A49097) that was built  
 CC using the pieces. The HIV DNA vaccine may be used as a prophylactic  
 CC vaccine and as a therapeutic vaccine in HIV infected patients.  
 CC  
 XX SQ Sequence 1038 BP; 191 A; 347 C; 331 G; 169 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 157  
 Score: 48.00 Length: 1038  
 Percent Similarity: 100.00% Matches: 7  
 Best Local Similarity: 100.00% Conservative: 0  
 Query Match: 100.00% Mismatches: 0  
 DB: 21 Indels: 0  
 Gaps: 0  
 US-09-877-606-7 (1-7) x AAA49097 (1-1038)  
 OY 1 LeuTPAsnTRppheasnile 7  
 ID 472 TTGTGGAACGTGTCACATC 492  
 Db  
 RESULT 15  
 AA062660  
 ID AA062660 standard; DNA; 1199 BP.  
 AC  
 XX AA062660;  
 XX  
 DT 06-JAN-1995 (first entry)  
 XX

DE gp41 HIV envelope protein coding sequence.  
 XX  
 KW Promoter; expression; lac; hepatitis; virus; protein; core; HCV;  
 KW CKS; CTP:CMP-3-deoxy-D-manno-oculosonate cytidyl transferase;  
 KW CMP-KDO synthetase; HIV; human immunodeficiency virus;  
 KW glycoprotein; envelope protein; gp41; ds.  
 XX  
 OS Human immunodeficiency virus.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT 1..1182  
 FT CDS /\*tag= a  
 FT /product= p41 HIV envelope protein.  
 XX  
 PN US5312737-A.  
 PD 17-MAY-1994.  
 PD  
 PF 11-MAR-1988; 88US-0167067.  
 PF  
 PR 11-MAR-1988; 88US-0167067.  
 PR 23-NOV-1988; 88US-0276263.  
 PR 24-AUG-1990; 90US-0573103.  
 PR 14-FEB-1992; 92US-0835878.  
 XX  
 PA (ABBO ) ABBOTT LAB.  
 PA  
 PI Bolling TJ, Mandeck W;  
 DR WPI: 1994-159120/19.  
 DR P-PSDB: AAR52687.  
 XX  
 PT Prodn. of recombinant hepatitis virus core protein - as fusion  
 PT protein with E. coli enzyme CKS  
 PT  
 PS Example 5; Figure 9; 63pp; English.  
 XX  
 CC The HIV gp41 envelope protein coding sequence was used in the  
 CC construction of numerous different expression vectors. The vectors  
 CC comprised (in a 5' to 3' order): a prokaryotic promoter; a  
 CC prokaryotic ribosome binding site; a gene fragment encoding CKS  
 CC (CTP:CMP-3-deoxy-D-manno- oculosonate cytidyl transferase or  
 CC CMP-KDO synthetase) protein and a gene fragment encoding at least  
 CC part of the HIV gp41 envelope protein. Plasmid pTR310B  
 CC comprises bases 1-14, 139-609 and 724-1199 of this sequence.  
 CC Plasmid pTR310A is identical except that it has a adenine base  
 CC deleted at position 813. Plasmid pTR321 comprises the whole sequence  
 CC minus bases 610-720 (hydrophobic region of gp140) and plasmid pTR322  
 CC comprises the whole sequence minus bases 15-143 (carboxy terminus of  
 CC gp120).  
 CC  
 XX SQ Sequence 1199 BP; 247 A; 301 C; 321 G; 330 T; 0 other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 183  
 Score: 48.00 Length: 1199  
 Percent Similarity: 100.00% Matches: 7  
 Best Local Similarity: 100.00% Conservative: 0  
 Query Match: 100.00% Mismatches: 0  
 DB: 15 Indels: 0  
 Gaps: 0  
 US-09-877-606-7 (1-7) x AA062660 (1-1199)  
 OY 1 LeuTPAsnTRppheasnile 7  
 ID 616 CTGTGGAACGTGTTAAACATC 636  
 Db  
 Search completed: March 25, 2003, 13:11:40  
 Job time : 28.3103 secs

Pred. No.: 128 Length: 862  
 Score: 48.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 17 Gaps: 0

US-09-877-606-7 (1-7) x AAT28647 (1-862)

OY 1 LeuTTPASnTRpPheAsnIle 7  
 |||||  
 Db 610 TTGTGGAATTGCTTTAACATA 630

RESULT 12

AAQ03968 ID AAQ03968 standard; DNA; 1026 BP.

AC AAO03968;

DT 04-SEP-1990 (first entry)

DE PENV9 encoding a recombinant peptide with HIV antigenicity.

XX human immunodeficiency virus; antigenicity; gp120; gp41;

KM poliovirus;

XX synthetic.

PN US4861707-A.

PD 29-AUG-1989.

PF 02-FEB-1987; 87US-0010056.

PR 02-FEB-1987; 87US-0010056.

PA (DUPO ) DU PONT DE NEMOURS & CO E I.

PI Ivanoff LA, Petteway SR;

DR WPI: 1990-122902/16.

DR P-PSDB; AAR04019.

PT Recombinant peptide with antigenicity of human immunodeficiency virus  
 comprises forty-six amino acids from gp120 C terminal and two hundred and  
 PT forty amino acids from gp41 N terminal

PS Claim 1; Fig 4; 13pp; English.

CC The sequence encodes ENV9, a recombinant peptide. ENV9 consists of about  
 CC 54 amino acids from the N-terminal of a poliovirus sequence (PEXC), as  
 CC well as those regions of HIV gp120 and gp41 detailed above.

CC ENV9 can be used to detect HIV antibodies. "False positive" results are  
 CC minimised as it contains few non-critical antigens. It can also be used  
 CC in a vaccine against HIV.

CC

XX Sequence 1026 BP; 335 A; 178 C; 275 G; 238 T; 0 other;

Alignment Scores:

Pred. No.: 155 Length: 1026  
 Score: 48.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 11 Gaps: 0

US-09-877-606-7 (1-7) x AAQ03968 (1-1026)

OY 1 LeuTTPASnTRpPheAsnIle 7  
 |||||

Db 775 TTGTGGAATTGCTTTAACATA 795

RESULT 13

AAQ28095 ID AAQ28095 standard; DNA; 1026 BP.

AC AAQ28095;

DT 12-FEB-1993 (first entry)

DE Recombinant DNA encoding ENV9.

XX Poliovirus; HIV; human immunodeficiency virus; AIDS; gp120; gp41; ss.

XX Synthetic.

PN US5141867-A.

PD 25-AUG-1992.

PF 02-FEB-1987; 87US-0010056.

PR 02-FEB-1987; 87US-0010056.

PA (DUPO ) DU PONT DE NEMOURS & CO E I.

PI Ivanoff LA, Petteway SR;

DR WPI: 1992-307861/37.

DR P-PSDB; AAR26787.

PT Nucleotide sequence encoding a HIV antigen - comprising the  
 C-terminal region of Gp120 and N-terminal of Gp41. for use in  
 PT diagnosis and in vaccines

PS Claim 3; Fig 4; 12pp; English.

CC The plasmid PENV9 was constructed by ligating the BglII/BamHI  
 CC fragment (nt 7196-8053) of HIV clone lambda BH10 with the BglII/  
 CC BamHI vector fragment of pEXC. The prod. was used to transform E.  
 CC coli strain HB101 and MM294. The transformants produced a protein  
 CC ENV9 contg. 349 amino acids corresp. to 54 amino acids from the N-  
 CC terminal of the poliovirus sequence of pEXC, 46 amino acids of the C-  
 CC terminal of the gp120 domain and 240 amino acids of the N-terminal  
 CC of the gp41 domain. The peptide comprises the early detection  
 CC antigen from gp120 and a highly reactive antigen from gp41 thus  
 CC providing broad sensitivity to HIV antibodies. The sensitivity of  
 CC the peptides permits the detection of HIV infection using small ants.

CC of diagnostic reagents. The peptides have improved stability as  
 CC compared to larger antigenic segments and contain fewer noncritical  
 CC antigens which minimises the possibility of false positive results  
 CC in detection of HIV infection. The peptides can also be used in  
 CC vaccines protective against HIV.

XX

XX Sequence 1026 BP; 335 A; 178 C; 275 G; 238 T; 0 other;

Alignment Scores:

Pred. No.: 155 Length: 1026  
 Score: 48.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 13 Gaps: 0

US-09-877-606-7 (1-7) x AAQ28095 (1-1026)

OY 1 LeuTTPASnTRpPheAsnIle 7  
 |||||

Db 775 TTGTGGAATTGCTTTAACATA 795

RESULT 14

AA449097 ID AAA449097 standard; DNA; 1038 BP.

AC AAA449097;



CC for each derivative of an early, primary HIV envelope gene. The first  
 CC stage in the production of the construct was the cloning of an HIV  
 CC envelope gene. A nucleotide sequence encoding this gene was then  
 CC created using codons from highly expressed mammalian genes. In the  
 CC next stage, snuts (AAA49060-A49079) were created by redesigning this  
 CC nucleotide construct so that restriction enzyme sites surrounded  
 CC functional regions of the sequence. The present sequence is one of  
 CC the pieces (AAA49080-A49092) that were created by assembling the snuts.  
 CC Each derivative of the envelope gene (AAA49093-A49097) was then built  
 CC using the pieces. The HIV DNA vaccine may be used as a prophylactic  
 CC vaccine and as a therapeutic vaccine in HIV infected patients.

SQ Sequence 800 BP; 163 A; 257 C; 255 G; 125 T; 0 other;

#### Alignment Scores:

Pred. No.:	118	Length:	800
Score:	48.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0

US-09-877-606-7 (1-7) x AAA49091 (1-800)

OY 1 LeuTrpAsnTrpPheAsnIle 7

Db 626 TTGTGGAACCTGCTCAACATC 646

#### RESULT 10

AAAN71256  
 ID AAAN71256 standard; DNA; 854 BP.

XX AAAN71256;

DT 29-MAY-1991 (first entry)

DE Sequence of HTLV-III provirus nucleotides 7199-8052, BHB-ENV,

DE clone G.

XX HIV; AIDS; ARC; vaccine; diagnosis; immunosassay; ss.

OS HTLV-III.

PN WC8704728-A.

PD 13-AUG-1987.

PF 03-FEB-1987; 87WO-US00225.

PR 03-FEB-1987; 87WO-US00225.

PR 25-SEP-1986; 86US-0911455.

PR 03-FEB-1986; 86US-0825597.

XX (CAMP-) CAMBRIDGE BIOSCIENC.

XX Beltz GA, Thorn RM, Marciani DJ, Hung CH;

XX WPI; 1987-235361/33.

XX Peptide fragments immunoreactive with HTLV-III antibodies - obtd.

PT by expressing cloned fragmented HTLV-III provirus nucleotide

PT sequences

XX Disclosure: Fig 19; 85pp; English.

XX The patent claims a peptide fragment which is immunoreactive to

CC antibodies against the HTLV-III virus and encoded for by the

CC HTLV-III provirus nucleotide sequences selected from nucleotides 653

CC to 1218, 2600 to 3911, 2743 to 4211, 6619 to 7198, and 7199 to 8052.

CC The peptide fragments are specifically recognised by antibodies

CC produced in response to the HTLV-III virus and can be used in

CC diagnostic assays. They can also be used to raise antibodies which

CC can be used in immunoassays and can also be used as a vaccine

CC against the HTLV-III virus.

SQ Sequence 854 BP; 286 A; 141 C; 229 G; 198 T; 0 other;

#### Alignment Scores:

Pred. No.:	127	Length:	854
Score:	48.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	8	Gaps:	0

US-09-877-606-7 (1-7) x AAAN71256 (1-854)

OY 1 LeuTrpAsnTrpPheAsnIle 7

Db 608 TTGTGGAACCTGCTTAAACATA 628

#### RESULT 11

AAT28647  
 ID AAT28647 standard; DNA; 862 BP.

XX AAT28647;

DT 27-NOV-1996 (first entry)

DE DNA fragment HE 117, encoding hybrid polypeptide E 117.

XX DNA fragment; HE 117; hybrid; fusion; polypeptide; protein; E 177;

XX human immunodeficiency virus; type 1; HIV-1; env; gene product;

XX serodiagnosis; plasmid pHE 117; ss.

XX Synthetic.

XX Key Location/Qualifiers

FT mat\_peptide 2..862

FT /\*tag= "a adenine 227 or 228, thymine 349 and

FT /note= "a adenine 658 have to be deleted before

XX RU2043413-C1.

XX 10-SEP-1995.

PF 28-FEB-1992; 92RU-0029952.

PR 28-FEB-1992; 92SU-4029952.

XX (BIOS-) BIOSERVIS BIOTECHN CO.

XX Sukhanova LL, Zaitsev IZ, Zvonarev AVU;

XX WPI; 1996-207765/21.

XX P-PSDB; AAR91826.

XX Hybrid polypeptide E 117 for detecting antibodies to HIV-1

PT comprises HIV-1 env protein fused to E.coli beta-galactosidase

PT Claim 2; Columns 5-8; 5pp; Russian.

XX The present sequence, the DNA fragment HE 117, encodes the hybrid

CC polypeptide E 177, which comprises the prod. of the HIV-1 env gene

CC and E. coli beta-galactosidase. The hybrid polypeptide is useful

CC for the immunodetection of HIV-1 antibodies, and therefore AIDS

CC serodiagnosis. E 177 is synthesised by the E. coli strain

CC VAPM NB-5873, obtd. by transforming E. coli with the plasmid

CC pHE 117, which carries the HIV-1 env gene fragment.

SQ Sequence 862 BP; 284 A; 141 C; 229 G; 208 T; 0 other;

Alignment Scores:

PT New recombinant HIV protein, DEV-1 - used for the specific, early  
PT detection of HIV antibodies in human biological samples  
XX  
PS Claim 6, Page 29; 37pp; English.  
XX  
CC The DEV-1 protein comprises the C-terminus of HIV gp120 and most of  
CC gp41. Two E.coli cytotoxic sections are deleted from this  
CC gp120-gp41 fusion protein to enable expression in E.coli cells.  
CC The DEV-1 protein is useful for the detection of HIV antibodies in  
CC human biological samples. The DEV-1 protein is highly  
CC immunoreactive with anti-HIV antibodies and can provide for early  
CC detection of HIV infection (seroconversion). The protein is smaller  
CC in size compared to prior art recombinant proteins, thus minimising  
CC the possibility of false positives. The DEV-1 protein tolerates a  
CC wide range of pH and buffer conditions and can be easily lyophilised  
CC then reconstituted into soln. In addn., the protein shows excellent  
CC stability over a wide range of temps. for long periods of time.  
XX  
SQ Sequence 699 BP; 247 A; 115 C; 184 G; 153 T; 0 other;  
XX  
Alignment Scores:  
Pred. No.: 102 Length: 699  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0  
US-09-877-606-7 (1-7) x AAT26993 (1-699)  
QY 1 LeuTrpAsnTrpPheAsnIle 7  
Db 523 TTGTGGAATTGGTTACACATA 543  
RESULT 8  
ID AAA49090 standard; DNA: 798 BP.  
XX  
XX AAA49090;  
XX  
DT 16-NOV-2000 (first entry)  
XX  
DE Piece 89p160 DNA for use in HIV DNA vaccine.  
XX  
XX HIV; human immunodeficiency virus; vaccine; AIDS; snut;  
KM silent nucleotide substitution; ds.  
XX  
OS Human immunodeficiency virus type 1.  
OS Synthetic.  
XX  
PN WO200029561-A2.  
XX  
PD 25-MAY-2000.  
XX  
PF 27-MAR-2000; 2000MO-DK00144.  
XX  
PR 29-MAR-1999; 99DK-0000427.  
PR 09-APR-1999; 99US-0128558.  
XX  
PA (STAT-) STATENS SERUM INST.  
XX  
PI Fomsgaard A;  
XX  
DR WPI: 2000-387778/33.  
DR P-PSDB; AAY99921.  
XX  
XX Producing nucleotide sequence construct with optimized codons for human  
PT immunodeficiency virus (HIV) genetic vaccine involves obtaining a first  
PT nucleotide sequence from a HIV patient, redesigning and assembling it  
PT with snuts -  
XX  
XX Example 4; Page 124-125; 150pp; English.  
XX

CC The present invention relates to a nucleotide construct with optimised  
CC codons for use as a human immunodeficiency virus (HIV) DNA vaccine. The  
CC construct uses codons from highly expressed mammalian proteins to code  
CC for each derivative of an early, primary HIV envelope gene. The first  
CC stage in the production of the construct was the cloning of an HIV  
CC envelope gene. A nucleotide sequence encoding this gene was then  
CC created using codons from highly expressed mammalian genes. In the  
CC next stage, snuts (AAA49060-A49079) were created by redesigning this  
CC nucleotide construct so that restriction enzyme sites surrounded  
CC functional regions of the sequence. The present sequence is one of  
CC the pieces (AAA49080-A49092) that were created by assembling the snuts.  
CC Each derivative of the envelope gene (AAA49093-A49097) was then built  
CC using the pieces. The HIV DNA vaccine may be used as a prophylactic  
CC vaccine and as a therapeutic vaccine in HIV infected patients.  
XX  
SQ Sequence 798 BP; 161 A; 260 C; 254 G; 123 T; 0 other;  
XX  
Alignment Scores:  
Pred. No.: 118 Length: 798  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 21 Gaps: 0  
US-09-877-606-7 (1-7) x AAA49090 (1-798)  
QY 1 LeuTrpAsnTrpPheAsnIle 7  
Db 626 TTGTGGAACGTGTTCAACATC 646  
RESULT 9  
ID AAA49091 standard; DNA: 800 BP.  
XX  
XX AAA49091;  
XX  
DT 16-NOV-2000 (first entry)  
XX  
DE Piece 89p150 DNA for use in HIV DNA vaccine.  
XX  
XX HIV; human immunodeficiency virus; vaccine; AIDS; snut;  
KM silent nucleotide substitution; ds.  
XX  
OS Human immunodeficiency virus type 1.  
OS Synthetic.  
XX  
PN WO200029561-A2.  
XX  
PD 25-MAY-2000.  
XX  
PF 27-MAR-2000; 2000MO-DK00144.  
XX  
PR 29-MAR-1999; 99DK-0000427.  
PR 09-APR-1999; 99US-0128558.  
XX  
PA (STAT-) STATENS SERUM INST.  
XX  
PI Fomsgaard A;  
XX  
DR WPI: 2000-387778/33.  
DR P-PSDB; AAY99922.  
XX  
XX Producing nucleotide sequence construct with optimized codons for human  
PT immunodeficiency virus (HIV) genetic vaccine involves obtaining a first  
PT nucleotide sequence from a HIV patient, redesigning and assembling it  
PT with snuts -  
XX  
XX Example 4; Page 126-127; 150pp; English.  
XX  
XX The present invention relates to a nucleotide construct with optimised  
CC codons for use as a human immunodeficiency virus (HIV) DNA vaccine. The  
CC construct uses codons from highly expressed mammalian proteins to code

PT	Producing nucleotide sequence construct with optimized codons for human immunodeficiency virus (HIV) genetic vaccine involves obtaining a first nucleotide sequence from a HIV patient, redesigning and assembling it with snuts -
PS	
PX	
XX	Claim 20; Page 119-120; 150pp; English.
CC	The present invention relates to a nucleotide construct with optimised codons for use as a human immunodeficiency virus (HIV) DNA vaccine. The construct uses codons from highly expressed mammalian proteins to code for each derivative of an early, primary HIV envelope gene. An HIV stage in the production of the construct was the cloning of an HIV envelope gene. A nucleotide sequence encoding this gene was then created using codons from highly expressed mammalian genes. In the next stage, snuts (AAA49060-A49079) were created by redesigning this nucleotide construct so that restriction enzyme sites surrounded functional regions of the sequence. The present sequence is one of the pieces (AAA49080-A49092) that were created by assembling the snuts. Each derivative of the envelope gene (AAA49093-A49097) was then built using the pieces. The HIV DNA vaccine may be used as a prophylactic vaccine and as a therapeutic vaccine in HIV infected patients.
CC	
CC	
CC	
CC	
SO	Sequence 432 BP; 96 A; 132 C; 129 G; 75 T; 0 other;
Alignment Scores:	
Pred. NO.:	60.7
Score:	48.00
Percent Similarity:	100.00%
Best local Similarity:	100.00%
Query Match:	100.00%
DB:	21
Gaps:	0
US-09-877-606-7 (1-7) x AAA49086 (1-432)	
OY	1 LeutrpAsnTrpPheAsnIle 7 
Db	260 TTGTGGAACCTGGTTCAACATC 280
RESULT 6	
AAAA9087	
ID	AAA49087 standard; DNA; 434 BP.
XX	
XX	AAA49087;
DT	16-NOV-2000 (first entry)
DE	Piece 49p150 DNA for use in HIV DNA vaccine.
KM	HIV; human immunodeficiency virus; vaccine; AIDS; snut;
KW	silent nucleotide substitution; ds.
XX	
OS	Human immunodeficiency virus type 1.
OS	Synthetic.
XX	
PN	WO200029561-A2.
PD	25-MAY-2000.
PF	27-MAR-2000; 2000WO-DK00144.
PR	29-MAR-1999; 99DK-0000427.
PR	09-APR-1999; 99US-0128558.
XX	
PA	(STAT-) STATENS SERUM INST.
XX	
F1	Fomsgaard A;
XX	
DR	WIPI; 2000-387778/33.
DR	P-PSDB; AA99918.
XX	
PT	Producing nucleotide sequence construct with optimized codons for human immunodeficiency virus (HIV) genetic vaccine involves obtaining a first nucleotide sequence from a HIV patient, redesigning and assembling it

PT	with snuts -		
XX	Claim 20; Page 121; 150pp; English.		
XX			
CC	The present invention relates to a nucleotide construct with optimised		
CC	codons for use as a human immunodeficiency virus (HIV) DNA vaccine. The		
CC	construct uses codons from highly expressed mammalian proteins to code		
CC	for each derivative of an early, primary HIV envelope gene. The first		
CC	stage in the production of the construct was the cloning of an HIV		
CC	envelope gene. A nucleotide sequence encoding this gene was then		
CC	created using codons from highly expressed mammalian genes. In the		
CC	next stage, snuts (AAA49060-A49079) were created by redesigning this		
CC	nucleotide construct so that restriction enzyme sites surrounded		
CC	functional regions of the sequence. The present sequence is one of		
CC	the pieces (AAA49080-A49092) that were created by assembling the snuts		
CC	Each derivative of the envelope gene (AAA49093-A49097) was then built		
CC	using the pieces. The HIV DNA vaccine may be used as a prophylactic		
CC	vaccine and as a therapeutic vaccine in HIV infected patients.		
XX			
XX	Sequence 434 BP; 98 A; 128 C; 131 G; 77 T; 0 other;		
Alignment Scores:			
Pred. No.:	61	Length:	434
Score:	48.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	21	Gaps:	0
US-09-877-606-7 (1-7) x AAA49087 (1-434)			
QY	1 LeuTRPAsnTRPheAsnIle 7		
Db			
	260 TTCTGGAGACTGCTTCACATC 280		
RESULT 7			
AAT26993	AAT26993 standard; cDNA; 699 BP.		
XX	AAT26993:		
XX	15-OCT-1996 (first entry)		
DE	Recombinant HIV protein DEV-1.		
XX	HIV; DEV-1; gp120; gp41; cytotoxic; detection; antibody;		
KW	seroconversion; infection; immunoreactive; ss.		
KM			
XX	Human immunodeficiency virus.		
OS			
XX	Key	Location/Qualifiers	
FH	misc_difference 229..231	/tag= a	
FT		/trans_except= 229..231, aa:Ala	
FT	misc_difference 613..615	/tag= b	
FT		/trans_except= 613..615; aa:Ile	
XX			
XX	WO9612023-A1.		
XX	25-APR-1996.		
XX			
XX	11-OCT-1995; 95MO-US13335.		
XX			
XX	17-OCT-1994; 94IL-0111311.		
XX			
XX	(DEVA-) DEVARON INC.		
XX			
XX	Devash Y;		
XX			
XX	WPI, 1996-222009/22.		
XX			
XX	P-PSDB; AAR93119.		
XX			

XX Example 11; Page 57; 82pp; English.  
 XX  
 CC The present specification describes a synthetic DNA sequence encoding  
 CC human immunodeficiency virus (HIV) env protein, or a fragment,  
 CC comprising codons optimised for expression in a mammalian host. The  
 CC present sequence represents a synthetic gene segment from an example of  
 CC the present invention. The new synthetic DNA sequence induces anti-HIV  
 CC neutralising antibody, HIV specific T cell immune responses, or  
 CC protective immune responses upon introduction into vertebrate tissue.  
 CC Including human tissue in vivo, where the DNA comprises a gene encoding  
 CC a HIV gag, gag protease or env gene product. The DNA can also be used  
 CC to induce an immune response against infection or disease caused by  
 CC virulent strains of HIV, optionally in conjunction with attenuated HIV,  
 CC killed HIV or HIV env, gag or pol protein, or to induce an anti-HIV  
 CC response in a primate when administered parenterally in conjunction  
 CC with interleukin 12. The new DNA can also be used to induce an antigen  
 CC presenting cell to stimulate cytotoxic and helper T cell proliferation  
 CC and effector functions, including lymphokine secretion specific to HIV  
 CC antigens, by expressing vertebrate cells to it in vivo, or as a vaccine  
 CC against HIV infection. The codons of new synthetic DNA are designed so  
 CC as to use the codons preferred by the projected host cell.  
 XX  
 SQ Sequence 387 BP; 91 A; 102 C; 113 G; 81 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 53.9 Length: 387  
 Score: 48.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 DB: 18  
 US-09-877-606-7 (1-7) x AAT85540 (1-387)  
 QY 1 LeuTRPAsnTRPheAsnIle 7  
 Db 230 CTGTGGAACGTGTCACACATC 250  
 RESULT 4  
 AAV21766  
 ID AAV21766 standard; DNA: 387 BP.  
 AC AAV21766;  
 XX  
 DT 20-JUL-1998 (first entry)  
 XX  
 DE HIV gp32 (opt) gene segment from Avrii site.  
 XX  
 KW VJuno; HIV vaccine; codon usage optimisation; HIV; immune response;  
 KW human immunodeficiency virus; hepatitis; tuberculosis; tumour;  
 KW env protein; gp32; recombinant; ss.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus.  
 XX  
 PN WO9748370-A2.  
 XX  
 PD 24-DEC-1997.  
 XX  
 PF 17-JUN-1997; 97WO-US10517.  
 XX  
 PR 16-JUL-1996; 96GB-0014943.  
 PR 21-JUN-1996; 96US-0020165.  
 PR 21-JUN-1996; 96US-0020166.  
 PR 16-JUL-1996; 96GB-0014942.  
 XX  
 PA (MERI ) MERCK & CO INC.  
 XX  
 PI Davies ME, Freed DC, Liu MA, Perry HC, Shiver JW;  
 XX WPI; 1998-062825/06.  
 DR  
 XX

PT DNA encoding protein with codon usage optimised for intended host  
 PT cell - specifically for DNA vaccines against human immune deficiency  
 PT virus, allows rev-independent expression of HIV genes  
 XX  
 XX Example 11; Pages 54-55; 113pp; English.  
 XX  
 CC This is a HIV gp32 gene segment from the Avrii site to the end of gp143  
 CC having optimal codon usage (opt) for expression. This can be used in an  
 CC expression vector for vaccine production. The expression vector contains  
 CC a new synthetic DNA, encoding a HIV env protein or its fragment and has  
 CC codons optimised for expression in a non-homologous host. The synthetic  
 CC DNA sequences are used to increase production of recombinant protein and  
 CC to induce anti-HIV neutralising antibody, HIV-specific T-cell immune  
 CC responses or protective immune responses in vertebrates, specifically as  
 CC HIV vaccines. Optimisation of codons results in increased expression of  
 CC the DNA in the host. The DNA induces antigen-presenting cells to  
 CC stimulate cytotoxic and helper T-cells, and effector functions such as  
 CC lymphokine secretion specific to HIV antigens. Cross-strain protection is  
 CC achieved without use of adjuvant and the synthetic DNA may provide a  
 CC longer-lasting response. The env gene includes a high proportion of  
 CC codons rarely used in mammals, so optimisation will allow this gene to  
 CC be expressed efficiently in absence of rev. The DNA sequences, more  
 CC generally, may express antigens from many other pathogens, e.g. hepatitis  
 CC or tuberculosis, and also from tumours.  
 XX  
 SQ Sequence 387 BP; 91 A; 102 C; 113 G; 81 T; 0 other;  
 Alignment Scores:  
 Pred. No.: 53.9 Length: 387  
 Score: 48.00 Matches: 7  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0  
 DB: 19  
 US-09-877-606-7 (1-7) x AAV21766 (1-387)  
 QY 1 LeuTRPAsnTRPheAsnIle 7  
 Db 230 CTGTGGAACGTGTCACACATC 250  
 RESULT 5  
 AAA49086  
 ID AAA49086 standard; DNA: 432 BP.  
 AC AAA49086;  
 XX  
 DT 16-NOV-2000 (first entry)  
 XX  
 DE Piece 49p160 DNA for use in HIV DNA vaccine.  
 XX  
 KW HIV; human immunodeficiency virus; vaccine; AIDS; snut;  
 KW silent nucleotide substitution; ds.  
 XX  
 OS Synthetic.  
 OS Human immunodeficiency virus type 1.  
 XX  
 PN WO200029561-A2.  
 XX  
 PD 25-MAY-2000.  
 XX  
 PF 27-MAR-2000; 2000WO-DK00144.  
 XX  
 PR 29-MAR-1999; 99DK-00000427.  
 PR 09-APR-1999; 99US-0128558.  
 XX  
 PA (STAT-) STATENS SERUM INST.  
 XX  
 PI Fomsgaard A;  
 XX WPI; 2000-387778/33.  
 DR P-PsDB; AAY99917.  
 DR  
 XX

```

XX 21-JUL-2000; 2000MO-US19816.
PF
XX 21-JUL-1999; 99US-0144965.
PR
XX (LEXI-) LEXIGEN PHARM CORP.
PA
XX Gallies SD, Lo KM, Wesolowski JS;
PI
XX WPI: 2001-182728/18.
DR
XX P-PSDB; AAB68575.
DR
XX Compositions for enhancing the immunogenicity of protein and peptide
PT antigens, comprise Fc fusion proteins, e.g. Fc-antigen fusion proteins
PT or Fc-adjutant fusion proteins -
XX
PS Example 1; Page 31; 78bp; English.
XX
XX The present invention relates to compositions for eliciting an enhanced
CC immune response against a preselected antigen. The compositions comprise
CC a fusion protein with an immunoglobulin heavy chain constant region
CC linked by a polypeptide to the antigen to elicit an immune response
CC against the antigen; or a nucleic acid sequence encoding the fusion
CC protein, where expression results in the production of the fusion
CC protein. The present sequence is a coding sequence for HIV gp41, which
CC was used in the present invention as an antigen sequence.
CC
SQ Sequence 148 BP; 39 A; 42 C; 41 G; 26 T; 0 other;
Alignment Scores:
Pred. No.: 19 Length: 148
Score: 48.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 22 Gaps: 0
US-09-877-606-7 (1-7) x AAF58385 (1-148)
QY 1 LeutrPAsnTrpPheAsnIle 7
DB 95 CTGTGGAACGTGTCACATC 115
RESULT 2
AAA49075
ID AAA49075 standard; DNA: 177 BP.
XX
AC AAA49075;
XX
DE 16-NOV-2000 (first entry)
XX
DE Snut 1890HindIII DNA used in HIV DNA vaccine.
XX
KW HIV; human immunodeficiency virus; vaccine; AIDS; snut;
KW silent nucleotide substitution; ds.
XX
OS Human immunodeficiency virus type 1.
XX
XX WO200029561-A2.
XX
XX 25-MAY-2000.
XX
XX 27-MAR-2000; 2000MO-DK00144.
XX
XX 29-MAR-1999; 99DK-0000427.
XX
XX 09-APR-1999; 99US-0128558.
XX
XX (STAT-) STATENS SERUM INST.
XX
XX Fomsgaard A;
XX
XX WPI: 2000-387778/33.
XX
XX P-PSDB; AAY99906.
XX

```

```

XX
XX Producing nucleotide sequence construct with optimized codons for human
PT immunodeficiency virus (HIV) genetic vaccine involves obtaining a first
PT nucleotide sequence from a HIV patient, redesigning and assembling it
PT with snuts -
XX
XX Example 3; Page 105; 150bp; English.
XX
XX The present invention relates to a nucleotide construct with optimised
CC codons for use as a human immunodeficiency virus (HIV) DNA vaccine. The
CC construct uses codons from highly expressed mammalian proteins to code
CC for each derivative of an early, primary HIV envelope gene. The first
CC stage in the production of the construct was the cloning of an HIV
CC envelope gene. A nucleotide sequence encoding this gene was then created
CC using codons from highly expressed mammalian genes. The present sequence
CC is one of the snuts (AAA49060-A49079) that were created by redesigning
CC this nucleotide construct so that restriction enzyme sites
CC surrounded functional regions of the sequence. The snuts were then
CC assembled into pieces (AAA49080-A49092). Each derivative of the envelope
CC gene (AAA49093-A49097) was then built using the pieces. The HIV DNA
CC vaccine may be used as a prophylactic vaccine and as a therapeutic
CC vaccine in HIV infected patients.
XX
SQ Sequence 177 BP; 31 A; 58 C; 50 G; 38 T; 0 other;
Alignment Scores:
Pred. No.: 23.1 Length: 177
Score: 48.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 21 Gaps: 0
US-09-877-606-7 (1-7) x AAA49075 (1-177)
QY 1 LeutrPAsnTrpPheAsnIle 7
DB 5 TTGTGGAACGTGTCACATC 25
RESULT 3
AAT85540
ID AAT85540 standard; DNA: 387 BP.
XX
AC AAT85540;
XX
DE 14-APR-1998 (first entry)
XX
DE gp32 (opt) synthetic gene segment.
XX
XX Human; tissue-specific plasminogen activator; tPA; leader; vaccine;
KW infection; HIV; env; human immunodeficiency virus; PCR primer; ss.
XX
OS Synthetic.
XX
OS Human immunodeficiency virus.
XX
XX WO9731115-A2.
XX
XX 28-AUG-1997.
XX
XX 18-FEB-1997; 97WO-US02294.
XX
XX 09-APR-1996; 96GB-0007293.
XX
XX 22-FEB-1996; 96US-0012082.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Davies M, Freed DC, Liu MA, Perry HC, Shiver JW;
XX
XX WPI: 1997-435167/40.
XX
XX Synthetic DNA encoding HIV env protein - comprises codon(s)
PT optimised for expression in mammalian host, useful as vaccine
PT against infection

```

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame-plus-p2n model

Run on: March 25, 2003, 11:26:42 ; Search time 27.3103 Seconds  
(Without alignments) 577.217 Million cell updates/sec

Title: US-09-877-606-7  
Perfect score: 48  
Sequence: 1 LNMWENI 7

Scoring table: BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 seqs, 112599159 residues  
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODE=frame+ p2n model -DEV=x1h  
-Q=/cgr21/USPTO.spool/US09877606/rnat\_18032003\_134800\_17645/app\_query.fasta.1.1059  
-DB=N-Geneseq.101002 -QF=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEC=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsun62 -TRANS=human40.cd1  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pcr -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09877606 @CNC 1.1.564 @runat.18032003\_134800\_17645 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEIOBERT -NRG\_SCORES=0 -WAIT -LONGLOG -DEF\_TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FCGAPOP=6 -FCGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N-Geneseq.101002.\*

- 1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*
- 2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*
- 3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*
- 4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*
- 5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*
- 6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*
- 7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*
- 8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*
- 9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*
- 10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*
- 11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*
- 12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*
- 13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*
- 14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*
- 15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*
- 16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*
- 17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*
- 18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*
- 19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*
- 20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*
- 21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*
- 22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*
- 23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*
- 24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	148	22 AAF58385	HIV gp41 coding se
2	48	100.0	177	21 AAA49075	Snut 1890HindIII D
3	48	100.0	387	18 AAT85540	gp32 (opt) syntnet
4	48	100.0	387	19 AAV21766	HIV gp32 (opt) 9e
5	48	100.0	432	21 AAA49086	piece 4gp160 DNA f
6	48	100.0	434	21 AAA49087	piece 4gp150 DNA f
7	48	100.0	699	17 AAT26933	Recombinant HIV pr
8	48	100.0	798	21 AAA49090	piece 8gp160 DNA f
9	48	100.0	800	21 AAA49091	piece 8gp150 DNA f
10	48	100.0	854	8 AAN71256	Sequence of HTLV-I
11	48	100.0	862	17 AAT28667	DNA fragment HE 11
12	48	100.0	1026	11 AAQ03968	PENY9 encoding a r
13	48	100.0	1026	13 AAQ28095	Recombinant DNA en
14	48	100.0	1038	21 AAA49097	synbX08- 41 gene d
15	48	100.0	1199	15 AAO62660	gp41 HIV envelope
16	48	100.0	1199	15 AAO55183	Sequence for a syn
17	48	100.0	1204	10 AAN0797	Synthetic p41 gene
18	48	100.0	1260	15 AAO66275	Sequence of HIV-1
19	48	100.0	1568	16 AAO80575	HIV-1 env gene. H
20	48	100.0	1568	17 AAT33295	HIV-1 env gene. H
21	48	100.0	2071	21 AAA49094	synbX08- 150 gene
22	48	100.0	2148	21 AAN97933	HIV-1 env DNA frag
23	48	100.0	2184	22 AAC82659	HIV-1/IIIB env c1o
24	48	100.0	2237	10 AAN90531	Sequence encoding
25	48	100.0	2263	10 AAN90398	HIV-1 env gene seg
26	48	100.0	2469	21 AAA49095	synbX08- 160 gene
27	48	100.0	2532	9 AAD29142	Human immunodefici
28	48	100.0	2553	24 AAN09949	HIV protein Htr6
29	48	100.0	2562	22 AAN20891	HIV gp41 DNA SEQ. I
30	48	100.0	2564	24 AAU41593	HIV-1 subtype C 1s
31	48	100.0	2565	22 AAF56366	HIV gene. Human 1
32	48	100.0	2607	7 AAN60077	DNA sequence of th
33	48	100.0	2641	13 AAO30970	Non-cleavable, sol
34	48	100.0	2694	19 AAV09364	GMCSF/HIV-1 chimet
35	48	100.0	2730	19 AAY36284	Human immunodefici
36	48	100.0	2945	16 AAT05127	HIV virus-1 gp160
37	48	100.0	3084	19 AAV09365	HIV-1/CD4 chimet
38	48	100.0	3156	7 AAN60128	Sequence of the en
39	48	100.0	3246	21 AAA48622	Lentivirus vector
40	48	100.0	4020	8 AAN71016	Sequence of LAV/HT
41	48	100.0	5703	24 ABR36429	HIV Cassette A1 DN
42	48	100.0	5703	24 ABR36432	HIV Cassette A2 DN
43	48	100.0	5925	24 ABR36431	HIV Cassette C1 DN
44	48	100.0	5925	24 ABR36434	HIV Cassette C2 DN
45	48	100.0	6229	21 AAA97934	HIV-1 env DNA frag

ALIGNMENTS

RESULT 1  
ID AAF58385 standard; DNA: 148 BP.  
XX AAF58385;  
AC AAF58385;  
XX 23-APR-2001 (first entry)  
DT 23-APR-2001 (first entry)  
XX HIV gp41 coding sequence.  
DE HIV gp41 coding sequence.  
XX Immunostimulant; vaccine; immune response; antibody; HIV; gp41;  
KW Immunoglobulin heavy chain constant region; ss.  
XX Human immunodeficiency virus type 1.  
OS Human immunodeficiency virus type 1.  
XX PN WO200107081-A1.  
XX PD 01-FEB-2001.



CDS  
1..>354  
/gene="env"  
/note="gp41 domain"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AG14336.1"  
/db\_xref="GI:10180893"  
/translation="AQQHLKLTVMGKIQALRYALERYLKDDQLGIGSCGKLLC  
TTTPVNNRTSWSEETMGNTLWMEWEROIDNTETIYKLLVESQNOQEKNEQDLALDQM  
ANLNNENNTNNMLYI"  
BASE COUNT 112 a 69 c 90 g 83 t  
ORIGIN

Alignment Scores:  
Pred. No.: 12.6 Length: 354  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-7 (1-7) x AF252127 (1-354)

RESULT 13  
LOCUS AF252128 354 bp DNA linear VRL 18-SEP-2000  
DEFINITION HIV-1 strain CAM70 subtype H from Cameroon envelope glycoprotein  
(env) gene, partial cds.  
ACCESSION AF252128  
VERSION AF252128.1 GI:10180894  
KEYWORDS Human immunodeficiency virus type 1.  
SOURCE Human immunodeficiency virus type 1  
ORGANISM Viruses; Retroviruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 354)  
Fonjongo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemji,G.A., Eno,L.T.,  
Mkengason,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
TITLE Presence of diverse human immunodeficiency virus type 1 viral  
variants in Cameroon  
JOURNAL AIDS Res. Hum. Retroviruses 16 (13), 1319-1324 (2000)  
MEDLINE 2041627  
PUBMED 10957729  
REFERENCE 2 (bases 1 to 354)  
Fonjongo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemji,G.A., Eno,L.T.,  
Mkengason,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2000) HIV/Retrovirus Disease Branch, DASTLR,  
Centers for Disease Control and Prevention, 1600 Clifton Road, Mail  
Stop G-19, Atlanta, GA 30333, USA  
FEATURES  
source  
1..354  
/organism="Human immunodeficiency virus type 1"  
/proviral  
/strain="CAM70"  
/db\_xref="taxon:11676"  
/country="Cameroon"  
/note="subtype: H"  
1..>354  
/gene="env"  
/note="gp41 domain"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AG14337.1"  
/db\_xref="GI:10180895"

/translation="AQQHMLQLTVMGKIQALRYALERYLKDDQLGIGSCGKLLC  
TTTPVNNRTSWSEETMGNTLWMEWEROIDNTETIYKLLVESQNOQEKNEQDLALDQM  
ASLNNENNTNNMLYI"  
BASE COUNT 114 a 69 c 93 g 78 t  
ORIGIN

Alignment Scores:  
Pred. No.: 12.6 Length: 354  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-7 (1-7) x AF252128 (1-354)

RESULT 14  
LOCUS AF405177 361 bp DNA linear VRL 22-JAN-2002  
DEFINITION HIV-1 isolate 972M117M from Zambia nonfunctional envelope  
glycoprotein (env) gene, partial sequence.  
ACCESSION AF405177  
VERSION AF405177.1 GI:18253916  
KEYWORDS Human immunodeficiency virus type 1.  
SOURCE Human immunodeficiency virus type 1  
ORGANISM Viruses; Retroviruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 361)  
Trask,S.A., Detdeyn,C.A., Fidel,I.U., Chen,Y., Meleth,S., Kasolo,F.,  
Musonda,R., Hunter,E., Gao,F., Allen,S. and Hahn,B.H.  
TITLE Molecular epidemiology of human immunodeficiency virus type 1  
transmission in a heterosexual cohort of discordant couples in  
Zambia  
J. Virol. 76 (1), 397-405 (2002)  
MEDLINE 21602569  
PUBMED 11739704  
REFERENCE 2 (bases 1 to 361)  
Gao,F.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Medicine, University of Alabama at  
Birmingham, 852 Kaul, 720 South 20th Street, Birmingham, AL 35294,  
USA  
FEATURES  
source  
1..361  
/organism="Human immunodeficiency virus type 1"  
/proviral  
/isolate="972M117M"  
/db\_xref="taxon:11676"  
/country="Zambia: Lusaka"  
/note="Isolated from a male patient in 1997  
subtype: C"  
1..>361  
/gene="env"  
/note="nonfunctional envelope glycoprotein due to  
mutation"

BASE COUNT 120 a 67 c 80 g 90 t 4 others  
ORIGIN

Alignment Scores:  
Pred. No.: 12.9 Length: 361  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0

US-09-877-606-7 (1-7) x AF405177 (1-361)



TITLE Direct Submission  
JOURNAL Submitted (04-APR-2000) HIV/Retrovirus Disease Branch, DASTLR,  
Centers for Disease Control and Prevention, 1600 Clifton Road, Mail  
Stop G-19, Atlanta, GA 30333, USA  
FEATURES  
Source 1. .354  
/organism="Human immunodeficiency virus type 1"  
/proviral  
/strain="CAM232"  
/db\_xref="taxon:11676"  
/country="Cameroon"  
/note="subtype: A"  
<1. .>354  
/gene="env"  
<1. .>354  
/gene="gag1 domain"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AG14328.1"  
/db\_xref="GI:10180877"  
/translation="AQQQLKLTVMGIGKQALVLAERYLRDQOLLGIGSGKLLC  
TTTPNNSSSNDIMNMWLMQDREISNTETITKLIQSQQEKNEDDLALDKM  
EGLMNNENTINMLMYI"  
BASE COUNT 120 a 67 c 85 g 82 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 12.6 Length: 354  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-877-606-7 (1-7) x AF252119 (1-354)  
QY 1 LeuTPasnrtpheasnlle 7  
|||||  
Db 313 CTGTGGAATTGGTTTACATTA 333  
RESULT 11  
AF252122 354 bp DNA linear VRL 18-SEP-2000  
LOCUS AF252122  
DEFINITION HIV-1 strain CAM256 subtype A from Cameroon envelope glycoprotein  
(env) gene, partial cds.  
ACCESSION AF252122  
VERSION AF252122.1 GI:10180882  
KEYWORDS  
SOURCE  
ORGANISM  
Human immunodeficiency virus type 1.  
Human immunodeficiency virus type 1  
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
1 (bases 1 to 354)  
Fonjungo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Mkenyasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
Presence of diverse human immunodeficiency virus type 1 viral  
variants in Cameroon  
AIDS Res. Hum. Retroviruses 16 (13), 1319-1324 (2000)  
JOURNAL  
MEDLINE 20414627  
PUBMED 10957729  
2 (bases 1 to 354)  
Fonjungo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Ngenasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
Direct Submission  
Submitted (04-APR-2000) HIV/Retrovirus Disease Branch, DASTLR,  
Centers for Disease Control and Prevention, 1600 Clifton Road, Mail  
Stop G-19, Atlanta, GA 30333, USA  
FEATURES  
Source 1. .354  
/organism="Human immunodeficiency virus type 1"

/proviral  
/strain="CAM256"  
/db\_xref="taxon:11676"  
/country="Cameroon"  
/note="subtype: A"  
<1. .>354  
/gene="env"  
<1. .>354  
/gene="gag1 domain"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AG14331.1"  
/db\_xref="GI:10180883"  
/translation="AQQHLKLTVMGIGKQALVLAERYLRDQOLLGIGSGKLLC  
TTTPNNSSSNDIMNMWLMQDREISNTETITKLIQSQQEKNEDDLALDKM  
ASLMNNENTINMLMYI"  
BASE COUNT 118 a 68 c 84 g 84 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 12.6 Length: 354  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-877-606-7 (1-7) x AF252122 (1-354)  
QY 1 LeuTPasnrtpheasnlle 7  
|||||  
Db 313 CTGTGGAATTGGTTTACATTA 333  
RESULT 12  
AF252127 354 bp DNA linear VRL 18-SEP-2000  
LOCUS AF252127  
DEFINITION HIV-1 strain CAM63 subtype A from Cameroon envelope glycoprotein  
(env) gene, partial cds.  
ACCESSION AF252127  
VERSION AF252127.1 GI:10180892  
KEYWORDS  
SOURCE  
ORGANISM  
Human immunodeficiency virus type 1.  
Human immunodeficiency virus type 1  
Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
1 (bases 1 to 354)  
Fonjungo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Mkenyasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
Presence of diverse human immunodeficiency virus type 1 viral  
variants in Cameroon  
AIDS Res. Hum. Retroviruses 16 (13), 1319-1324 (2000)  
JOURNAL  
MEDLINE 20414627  
PUBMED 10957729  
2 (bases 1 to 354)  
Fonjungo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Ngenasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
Direct Submission  
Submitted (04-APR-2000) HIV/Retrovirus Disease Branch, DASTLR,  
Centers for Disease Control and Prevention, 1600 Clifton Road, Mail  
Stop G-19, Atlanta, GA 30333, USA  
FEATURES  
Source 1. .354  
/organism="Human immunodeficiency virus type 1"  
/proviral  
/strain="CAM63"  
/db\_xref="taxon:11676"  
/country="Cameroon"  
/note="subtype: A"  
<1. .>354  
/gene="env"

REFERENCE 1 lentivirus group.  
AUTHORS 1 (bases 1 to 354)  
Fonjongo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Nkengasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
TITLE Presence of diverse human immunodeficiency virus type 1 viral  
variants in Cameroon  
JOURNAL AIDS Res. Hum. Retroviruses 16 (13), 1319-1324 (2000)  
MEDLINE 20414627  
PUBMED 10957729  
REFERENCE 2 (bases 1 to 354)  
Fonjongo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Nkengasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2000) HIV/Retrovirus Disease Branch, DASTLR,  
Centers for Disease Control and Prevention, 1600 Clifton Road, Mail  
Stop G-19, Atlanta, GA 30333, USA  
FEATURES  
source  
1. 354  
/organism="Human immunodeficiency virus type 1"  
/proviral  
/strain="CAM161"  
/db\_xref="taxon:11676"  
/country="Cameroon"  
/note="subtype: A"  
<1. .>354  
/gene="env"  
<1. .>354  
/gene="env"  
/note="gp41 domain"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AAG14322.1"  
/translation="AQQHLKLTWGIKQLQARVLALESYLRDQQLGIGWCSGKILC  
TTVPWNSSMSDDIMNMTWLDREVSYTDIIYHLIESONQOEKNEODLALDKW  
ASLWNFNTIKWLYI"  
BASE COUNT 115 a 70 c 83 g 86 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 12.6 Length: 354  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-877-606-7 (1-7) x AF252103 (1-354)  
QY 1 LeuTPasnrtrpPheasnlle 7  
DB 313 CTATGGAATTGTTTAAACATA 333  
RESULT 9  
AF252113 354 bp DNA linear VRL 18-SEP-2000  
LOCUS HIV-1 strain CAM203 subtype A from Cameroon envelope glycoprotein  
DEFINITION (env) gene, partial cds.  
ACCESSION AF252113  
VERSION AF252113.1 GI:10180864  
KEYWORDS Human immunodeficiency virus type 1.  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 354)  
Fonjongo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Nkengasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
TITLE Presence of diverse human immunodeficiency virus type 1 viral  
variants in Cameroon

JOURNAL AIDS Res. Hum. Retroviruses 16 (13), 1319-1324 (2000)  
MEDLINE 20414627  
PUBMED 10957729  
REFERENCE 2 (bases 1 to 354)  
Fonjongo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Nkengasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
TITLE Direct Submission  
JOURNAL Submitted (04-APR-2000) HIV/Retrovirus Disease Branch, DASTLR,  
Centers for Disease Control and Prevention, 1600 Clifton Road, Mail  
Stop G-19, Atlanta, GA 30333, USA  
FEATURES  
source  
1. 354  
/organism="Human immunodeficiency virus type 1"  
/proviral  
/strain="CAM203"  
/db\_xref="taxon:11676"  
/country="Cameroon"  
/note="subtype: A"  
<1. .>354  
/gene="env"  
<1. .>354  
/gene="env"  
/note="gp41 domain"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="AAG14322.1"  
/translation="AQQHLKLTWGIKQLQARVLALESYLRDQQLGIGWCSGKILC  
TTVPWNSSMSDDIMNMTWLDREVSYTDIIYHLIESONQOEKNEODLALDKW  
ASLWNFNTIKWLYI"  
BASE COUNT 121 a 65 c 87 g 81 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 12.6 Length: 354  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 14 Gaps: 0  
US-09-877-606-7 (1-7) x AF252113 (1-354)  
QY 1 LeuTPasnrtrpPheasnlle 7  
DB 313 CTGTGGAATTGTTTAAACATA 333  
RESULT 10  
AF252119 354 bp DNA linear VRL 18-SEP-2000  
LOCUS HIV-1 strain CAM232 subtype A from Cameroon envelope glycoprotein  
DEFINITION (env) gene, partial cds.  
ACCESSION AF252119  
VERSION AF252119.1 GI:10180876  
KEYWORDS Human immunodeficiency virus type 1.  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Viruses; Retroid viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 354)  
Fonjongo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Nkengasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.  
TITLE Presence of diverse human immunodeficiency virus type 1 viral  
variants in Cameroon  
JOURNAL AIDS Res. Hum. Retroviruses 16 (13), 1319-1324 (2000)  
MEDLINE 20414627  
PUBMED 10957729  
REFERENCE 2 (bases 1 to 354)  
Fonjongo,P.N., Mpoudi,E.N., Torimiro,J.N., Alemnji,G.A., Eno,L.T.,  
Nkengasong,J.N., Gao,F., Rayfield,M., Folks,T.M., Pieniazek,D. and  
Lal,R.B.



```

/note="gp41"
/codon_start=1
/product="envelope glycoprotein"
/db_xref="GI:4151426"
/translation="AQQHLKLTWGIKQLQARVLAERYLRDQQLGIMGSGKLLC
TTNVPWNSTWNSKNTYDDIWGNMTWLQWMDREINNYETIYRLIEESQNOQEKNDQDLIA
LDKWEGLMWNFNITN"

BASE COUNT      122 a      66 c      84 g      79 t
ORIGIN

Alignment Scores:
Pred. No.:      12.5      Length:      351
Score:          48.00     Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels: 0
DB:             14      Gaps: 0

US-09-877-606-7 (1-7) x AF096332 (1-351)

OY      1      LeuTnpasnrTppheasnile 7
Db      325      CTGTGGAATGTTTAACATA 345

RESULT 4
LOCUS    AF096339      351 bp      RNA      linear      VRL 13-JAN-1999
DEFINITION HIV-1 strain 96USSN07 from USA, envelope glycoprotein (env) gene,
partial cds.
ACCESSION AF096339
VERSION    AF096339.1      GI:4151439
KEYWORDS
SOURCE      Human immunodeficiency virus type 1.
ORGANISM    Human immunodeficiency virus type 1
            Viruses; Retrovirdae; Lentiviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1 (bases 1 to 351)
AUTHORS    Sullivan,P.J., Do,A., Ellenberger,D.L., Pau,C.-P., Paul,S.,
            Kalish,M., Robbins,K., Lal,R., Storch,C., Schable,C.A., Wise,H.,
            Tetteh,C., Jones,J. and Ward,J.
            Surveillance of Central African Nationals living in the United
            States reveals multiple subtypes of HIV-1 Group M and Group O
            unpublished
            2 (bases 1 to 351)
JOURNAL    Sullivan,P.J., Do,A., Ellenberger,D.L., Pau,C.-P., Paul,S.,
AUTHORS    Kalish,M., Robbins,K., Lal,R., Storch,C., Schable,C.A., Wise,H.,
            Tetteh,C., Jones,J. and Ward,J.
            Direct Submission
            Submitted (02-OCT-1998) HARB, CDC, 1600 Clifton Road, Atlanta, GA
            30333, USA

FEATURES
Source      1..351
            location/Qualifiers
            /organism="Human immunodeficiency virus type 1"
            /strain="96USSN07"
            /db_xref="taxon:11676"
            /country="USA"
            /note="subtype: B"
            <1..>351
            /gene="env"
            /gene="env"
            <1..>351
            /note="gp41"
            /codon_start=1
            /product="envelope glycoprotein"
            /protein_id="AAD04414.1"
            /db_xref="GI:4151440"
            /translation="AQQHMLQLTWGIKQLQARVLAERYLRDQQLGIMGSGKLLC
            TTSVPNANSMNSKSLDKIWNMTWMEWEIEIYNLTGLIYNLLEESQNOQEKNDQELLE
            FDKWALIMWPNITN"

BASE COUNT      124 a      62 c      81 g      83 t      1 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      12.5      Length:      351
Score:          48.00     Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels: 0
DB:             14      Gaps: 0

US-09-877-606-7 (1-7) x AF096339 (1-351)

OY      1      LeuTnpasnrTppheasnile 7
Db      325      TTGTGGAATGTTTAACATA 345

RESULT 5
LOCUS    AF096350      351 bp      DNA      linear      VRL 13-JAN-1999
DEFINITION HIV-1 strain 96USSN48 from USA, envelope glycoprotein (env) gene,
partial cds.
ACCESSION AF096350
VERSION    AF096350.1      GI:4151461
KEYWORDS
SOURCE      Human immunodeficiency virus type 1.
ORGANISM    Human immunodeficiency virus type 1
            Viruses; Retrovirdae; Lentiviridae; Lentivirus; Primate
            lentivirus group.
REFERENCE   1 (bases 1 to 351)
AUTHORS    Sullivan,P.J., Do,A., Ellenberger,D.L., Pau,C.-P., Paul,S.,
            Kalish,M., Robbins,K., Lal,R., Storch,C., Schable,C.A., Wise,H.,
            Tetteh,C., Jones,J. and Ward,J.
            Surveillance of Central African Nationals living in the United
            States reveals multiple subtypes of HIV-1 Group M and Group O
            unpublished
            2 (bases 1 to 351)
JOURNAL    Sullivan,P.J., Do,A., Ellenberger,D.L., Pau,C.-P., Paul,S.,
AUTHORS    Kalish,M., Robbins,K., Lal,R., Storch,C., Schable,C.A., Wise,H.,
            Tetteh,C., Jones,J. and Ward,J.
            Direct Submission
            Submitted (02-OCT-1998) HARB, CDC, 1600 Clifton Road, Atlanta, GA
            30333, USA

FEATURES
Source      1..351
            location/Qualifiers
            /organism="Human immunodeficiency virus type 1"
            /proviral
            /strain="96USSN48"
            /db_xref="taxon:11676"
            /country="USA"
            <1..>351
            /gene="env"
            /gene="env"
            <1..>351
            /note="gp41"
            /codon_start=1
            /product="envelope glycoprotein"
            /protein_id="AAD04425.1"
            /db_xref="GI:4151462"
            /translation="AQQHMLQLTWGIKQLQARVLAERYLRDQQLGIMGSGKLLC
            TTNVPWNSSMSKNSYEIWNMTWMEWEIENYTLIYTLIEAQYQQEKNDQDLIA
            LDTWASIMWPNITN"

BASE COUNT      117 a      67 c      87 g      80 t
ORIGIN

Alignment Scores:
Pred. No.:      12.5      Length:      351
Score:          48.00     Matches:      7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:    100.00% Indels: 0
DB:             14      Gaps: 0

US-09-877-606-7 (1-7) x AF096350 (1-351)

OY      1      LeuTnpasnrTppheasnile 7

```

AX028583  
LOCUS AX028583 177 bp DNA linear PAT 16-SEP-2000  
DEFINITION Sequence 31 from Patent W0029561.  
ACCESSION AX028583  
VERSION AX028583.1 GI:10189736  
KEYWORDS  
SOURCE Human immunodeficiency virus.  
ORGANISM Human immunodeficiency virus  
Viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 177)  
AUTHORS Fomsgaard, A.  
TITLE Method for producing a nucleotide construct with optimised codons for an hiv genetic vaccine based on a primary, early hiv isolate and synthetic envelope bx08 constructs  
Patent: WO 0029561-A 31 25-MAY-2000;  
JOURNAL STATENS SERUMINSTITUT (DK) : FOMSGAARD ANDERS (DK)  
FEATURES  
Source location/Qualifiers  
1..177  
/organism="Human immunodeficiency virus"  
/db\_xref="taxon:12721"  
CDS  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="GAC09219.1"  
/db\_xref="GI:10189737"  
/translation="KLELVQHHQLAVYHODFHHNGRPPRAHVRHRAEHPRAAG  
LQPELFPDPPARAAR"  
BASE COUNT 31 a 58 c 50 g 38 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 6.15 Length: 177  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-877-606-7 (1-7) x AX028583 (1-177)  
QY 1 LeutrpasanttrpheaSnile 7  
|||||  
Db 5 TTGTGGAACGTGTTACACATC 25  
RESULT 2  
AF096327 351 bp DNA linear VRL 13-JAN-1999  
LOCUS HIV-1 strain 96USNG31 from USA, envelope glycoprotein (env) gene,  
partial cds.  
AF096327  
AF096327.1 GI:4151415  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1  
Viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 351)  
AUTHORS Sullivan, P.J., Do, A., Ellenberger, D.L., Pau, C.-P., Paul, S.,  
Kaliish, M., Robbins, K., Lal, R., Storch, C., Schable, C.A., Wise, H.,  
Reteh, C., Jones, J. and Ward, J.  
TITLE Surveillance of Central African Nationals living in the United States reveals multiple subtypes of HIV-1 Group M and Group O  
Unpublished  
JOURNAL 2 (bases 1 to 351)  
REFERENCE Sullivan, P.J., Do, A., Ellenberger, D.L., Pau, C.-P., Paul, S.,  
AUTHORS Kaliish, M., Robbins, K., Lal, R., Storch, C., Schable, C.A., Wise, H.,  
Reteh, C., Jones, J. and Ward, J.  
TITLE Direct Submission  
JOURNAL Submitted (02-OCT-1998) HARB, CDC, 1600 Clifton Road, Atlanta, GA  
30333, USA  
FEATURES  
Source location/Qualifiers  
1..351

/organism="Human immunodeficiency virus type 1"  
/proviral  
/strain="96USNG31"  
/db\_xref="taxon:11676"  
/country="USA"  
/note="subtype: C"  
CDS  
/gene="env"  
/gene="env"  
/note="gp41"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="A0404402.1"  
/db\_xref="GI:4151416"  
/translation="AQQHMLQFTWGIKQARVLAIERLYKPDQLGIWCGSGKLLC  
TTTPVPMSSNSNSKSYKDIWEMTMQMQDREIDNTYITVNLSESQIQQRKNEEDLLA  
LDSNQLNNFNFSH"  
BASE COUNT 123 a 64 c 77 g 87 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 12.5 Length: 351  
Score: 48.00 Matches: 7  
Percent Similarity: 100.00% Conservative: 0  
Best local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: Gaps: 0  
US-09-877-606-7 (1-7) x AF096327 (1-351)  
QY 1 LeutrpasanttrpheaSnile 7  
|||||  
Db 325 CTGTGGAATGTTTACACATA 345  
RESULT 3  
AF096332 351 bp DNA linear VRL 13-JAN-1999  
LOCUS HIV-1 strain 96USCM11 from USA, envelope glycoprotein (env) gene,  
partial cds.  
AF096332  
AF096332.1 GI:4151425  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE Human immunodeficiency virus type 1.  
ORGANISM Human immunodeficiency virus type 1  
Viruses; Retroviridae; Lentivirus; Primate  
lentivirus group.  
REFERENCE 1 (bases 1 to 351)  
AUTHORS Sullivan, P.J., Do, A., Ellenberger, D.L., Pau, C.-P., Paul, S.,  
Kaliish, M., Robbins, K., Lal, R., Storch, C., Schable, C.A., Wise, H.,  
Reteh, C., Jones, J. and Ward, J.  
TITLE Surveillance of Central African Nationals living in the United States reveals multiple subtypes of HIV-1 Group M and Group O  
Unpublished  
JOURNAL 2 (bases 1 to 351)  
REFERENCE Sullivan, P.J., Do, A., Ellenberger, D.L., Pau, C.-P., Paul, S.,  
AUTHORS Kaliish, M., Robbins, K., Lal, R., Storch, C., Schable, C.A., Wise, H.,  
Reteh, C., Jones, J. and Ward, J.  
TITLE Direct Submission  
JOURNAL Submitted (02-OCT-1998) HARB, CDC, 1600 Clifton Road, Atlanta, GA  
30333, USA  
FEATURES  
Source location/Qualifiers  
1..351  
/organism="Human immunodeficiency virus type 1"  
/proviral  
/strain="96USCM11"  
/db\_xref="taxon:11676"  
/country="USA"  
/note="subtype: A"  
CDS  
/gene="env"  
/gene="env"  
/note="gp41"  
/codon\_start=1  
/product="envelope glycoprotein"  
/protein\_id="A0404402.1"  
/db\_xref="GI:4151416"  
/translation="AQQHMLQFTWGIKQARVLAIERLYKPDQLGIWCGSGKLLC  
TTTPVPMSSNSNSKSYKDIWEMTMQMQDREIDNTYITVNLSESQIQQRKNEEDLLA  
LDSNQLNNFNFSH"

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:35:13 ; Search time 326.517 Seconds  
(Without alignments)  
623.917 Million cell updates/sec

Title: US-09-877-606-7  
Sequence: 1 LMWPMNI 7

Scoring table: BLOSUM62  
Xgapop 10.0 , Ygapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n.model -DEV=xlh  
-Q/cg92\_1/USFTO.spool/US09877606/runatc\_18032003\_134800\_17655/app\_query.fasta.1.1059  
-DB-GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0  
-UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi -LIST=45  
-OUTFMN-200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-USFR=US09877606.ecgn.1.1.3852.gunat.18032003.134800.17655 -NCPU=6 -ICPU=3  
-NO\_XLPXY -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV-TIMEOUT=120  
-WARN-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELeXT=7

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vi:\*  
15: em\_ba:\*  
16: em\_fun:\*  
17: em\_hum:\*  
18: em\_in:\*  
19: em\_mu:\*  
20: em\_om:\*  
21: em\_or:\*  
22: em\_ov:\*  
23: em\_pac:\*  
24: em\_ph:\*  
25: em\_pl:\*  
26: em\_ro:\*  
27: em\_sts:\*  
28: em\_un:\*

29: em\_vl:\*  
30: em\_htg\_hum:\*  
31: em\_htg\_iny:\*  
32: em\_htg\_other:\*  
33: em\_htg\_mus:\*  
34: em\_htg\_pln:\*  
35: em\_htg\_rtd:\*  
36: em\_htg\_mam:\*  
37: em\_htg\_vrt:\*  
38: em\_sy:\*  
39: em\_htgo\_hum:\*  
40: em\_htgo\_mus:\*  
41: em\_htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed.  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	100.0	177	6 AX028583	AX028583 Sequence
2	48	100.0	351	14 AF096337	AF096337 HIV-1 str
3	48	100.0	351	14 AF096332	AF096332 HIV-1 str
4	48	100.0	351	14 AF096339	AF096339 HIV-1 str
5	48	100.0	351	14 AF096350	AF096350 HIV-1 str
6	48	100.0	354	14 AF252095	AF252095 HIV-1 str
7	48	100.0	354	14 AF252102	AF252102 HIV-1 str
8	48	100.0	354	14 AF252103	AF252103 HIV-1 str
9	48	100.0	354	14 AF252113	AF252113 HIV-1 str
10	48	100.0	354	14 AF252119	AF252119 HIV-1 str
11	48	100.0	354	14 AF252122	AF252122 HIV-1 str
12	48	100.0	354	14 AF252127	AF252127 HIV-1 str
13	48	100.0	354	14 AF252128	AF252128 HIV-1 str
14	48	100.0	361	14 AF405177	AF405177 HIV-1 str
15	48	100.0	365	14 AF405173	AF405173 HIV-1 str
16	48	100.0	366	14 AF034037	AF034037 HIV-1 str
17	48	100.0	366	14 AF034062	AF034062 HIV-1 str
18	48	100.0	366	14 AF034065	AF034065 HIV-1 str
19	48	100.0	366	14 AF034066	AF034066 HIV-1 str
20	48	100.0	366	14 AF034068	AF034068 HIV-1 str
21	48	100.0	366	14 AF034069	AF034069 HIV-1 str
22	48	100.0	366	14 AF034072	AF034072 HIV-1 str
23	48	100.0	366	14 AF034075	AF034075 HIV-1 str
24	48	100.0	366	14 AF113579	AF113579 HIV-1 str
25	48	100.0	366	14 AF165535	AF165535 HIV-1 str
26	48	100.0	366	14 AF165538	AF165538 HIV-1 str
27	48	100.0	366	14 AF165544	AF165544 HIV-1 str
28	48	100.0	366	14 AF165555	AF165555 HIV-1 str
29	48	100.0	366	14 AF190947	AF190947 HIV-1 str
30	48	100.0	366	14 AF190952	AF190952 HIV-1 str
31	48	100.0	366	14 AF190956	AF190956 HIV-1 str
32	48	100.0	366	14 AF206051	AF206051 HIV-1 str
33	48	100.0	366	14 AF206057	AF206057 HIV-1 str
34	48	100.0	366	14 AF206058	AF206058 HIV-1 str
35	48	100.0	366	14 AF220675	AF220675 HIV-1 str
36	48	100.0	366	14 AF220681	AF220681 HIV-1 str
37	48	100.0	366	14 AF220683	AF220683 HIV-1 str
38	48	100.0	366	14 AF220685	AF220685 HIV-1 str
39	48	100.0	366	14 AF220690	AF220690 HIV-1 str
40	48	100.0	366	14 AF230721	AF230721 HIV-1 str
41	48	100.0	366	14 AF307651	AF307651 HIV-1 str
42	48	100.0	366	14 AF307652	AF307652 HIV-1 str
43	48	100.0	366	14 AF307690	AF307690 HIV-1 str
44	48	100.0	366	14 AF307730	AF307730 HIV-1 str
45	48	100.0	366	14 AF343902	AF343902 HIV-1 str

RESULT 1

#### ALIGNMENTS

**THIS PAGE BLANK (1857)**

FILED DATE: 08-JAN-2001  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/376,184  
FILED DATE: 17-AUG-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/987,867  
FILED DATE: 09-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/389,459  
FILED DATE: 15-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/087,009  
FILED DATE: 01-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauro, Peter C.  
REGISTRATION NUMBER: 32,360  
REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-756-551A-8

Query Match 100.0%; Score 48; DB 10; Length 519;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LMNMFNI 7  
| | | | | | | |  
Db 465 LMNMFNI 471

Search completed: March 18, 2003, 14:52:57  
Job time: 1.62069 secs



Db 622 NNNTWMEWDREINNTYSLIHSLEESQNOQEKNEQELLELDKRWASLMMWFI 675

## RESULT 15

ID 090178 PRELIMINARY; PRT; 854 AA.

AC 090178; 01-NOV-1998 (TREMBlrel. 08, Created)

DT 01-NOV-1998 (TREMBlrel. 08, last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, last annotation update)

DE Envelope glycoprotein.

GN ENV.

OS Human immunodeficiency virus type 1.

OC Viruses; Retroid viruses; Retroviridae; Lentivirus.

OX NCBI\_Taxid-11676;

RN [1]

RP SEQUENCE FROM N.A. MEDLINE-95074930; PubMed-7983770;

RA Fang H., Pincus S.H.;

RT "Unique insertion sequence and pattern of CD4 expression in variants

RT selected with immunotoxins from human immunodeficiency virus type 1-

RT infected T cells."

RL J. Virol. 69:75-81(1995).

RN [2]

RP SEQUENCE FROM N.A.

RA Fang H., Pincus S.H.;

RT "Spontaneous activation of human immunodeficiency virus type 1 in an

RT immunotoxin-resistant variant T cell line."

RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.

DR EMBL: AF070521; AAC28452.1; -

DR InterPro: IPR000328; Env\_GP41.

DR InterPro: IPR000777; GP120.

DR Pfam: PF00516; GP120; 1.

DR Pfam: PF00517; GP41; 1.

DR AIDS: Coat protein; Glycoprotein; Polypotein; Transmembrane.

SO SEQUENCE 854 AA; 96885 MW; 0808F3AED27C693B CRC64;

Query Match 91.0%; Score 264; DB 15; Length 854;

Best Local Similarity 90.4%; Pred. No. 4, 3e-19;

Matches 47; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NNNTWMEWDREINNTYSLIHSLEESQNOQEKNEQELLELDKRWASLMMWFI 52

Db 622 NNNTWMEWDREINNTYSLIHSLEESQNOQEKNEQELLELDKRWASLMMWFI 673

Search completed: March 18, 2003, 14:36:09  
Job time : 30.202 secs

DR InterPro; IPR000777; GP120.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS: Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 854 AA; 97124 MW; ABA0A1BF36D5595A CRC64;

Query Match 92.4%; Score 268; DB 15; Length 854;  
 Best Local Similarity 92.3%; Pred. No. 1.7e-19;  
 Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTLEMDREINNTSLIHSLSIESQNOEKLEQELLELDKWSLMMWNI 52  
 I:|||||  
 DB 622 NMTWMEDEINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMWNI 673

RESULT 12

OY 072502 PRELIMINARY; PRT; 854 AA.  
 AC 072502;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Env polyprotein.  
 GN Env.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NL4-3;  
 RX MEDLINE=96036482; PubMed=7483282;  
 RA Salminen M.O., Koch C., Sanders-Buell E., Ehrenberg P.K.,  
 RA Michael N.L., Carr J.K., Burke D.S., McCutchan F.E.;  
 RT "Recovery of virtually full-length HIV-1 provirus of diverse subtypes  
 from primary virus cultures using the polymerase chain reaction.";  
 RL Virology 213:80-86(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NL4-3;  
 RX MEDLINE=86281827; PubMed=3016298;  
 RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,  
 RA Martin M.A.;  
 RT "Production of acquired immunodeficiency syndrome-associated  
 retrovirus in human and nonhuman cells transfected with an infectious  
 molecular clone.";  
 RL J. Virol. 59:284-291(1986).  
 DR EMBL; U26942; AAB60578.1;  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS: Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 FT CONFLICT 214 214 H -> L (IN REF. 2).  
 FT CONFLICT 530 530 A -> S (IN REF. 2).  
 FT CONFLICT 739 739 G -> D (IN REF. 2).  
 SQ SEQUENCE 854 AA; 97005 MW; FF2264B3841D1220 CRC64;

Query Match 92.4%; Score 268; DB 15; Length 854;  
 Best Local Similarity 92.3%; Pred. No. 1.7e-19;  
 Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTLEMDREINNTSLIHSLSIESQNOEKLEQELLELDKWSLMMWNI 52  
 I:|||||  
 DB 622 NMTWMEDEINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMWNI 673

RESULT 13

OY 074599 PRELIMINARY; PRT; 856 AA.  
 AC 074599;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE Env.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MCK1;  
 RX MEDLINE=90101366; PubMed=1688473;  
 RA Cloyd M.W., Moore B.E.;  
 RT "Spectrum of Biological Properties of Human Immunodeficiency Virus  
 (HIV-1) Isolates.";  
 RL Virology 174:103-116(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MCK1;  
 RA Iwataani Y.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D86068; BAA12995.1;  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS: Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 856 AA; 97287 MW; 238042A234C56685 CRC64;

Query Match 92.4%; Score 268; DB 15; Length 856;  
 Best Local Similarity 92.3%; Pred. No. 1.7e-19;  
 Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTLEMDREINNTSLIHSLSIESQNOEKLEQELLELDKWSLMMWNI 52  
 I:|||||  
 DB 624 NMTWMEDEINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMWNI 675

RESULT 14

OY 074090 PRELIMINARY; PRT; 856 AA.  
 AC 074090;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Env.  
 GN ENV.  
 OS Human immunodeficiency virus type 1.  
 OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxID=11676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PM213;  
 RX MEDLINE=90101366; PubMed=1688473;  
 RA Cloyd M.W., Moore B.E.;  
 RT "Spectrum of Biological Properties of Human Immunodeficiency Virus  
 (HIV-1) Isolates.";  
 RL Virology 174:103-116(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PM213;  
 RA Iwataani Y.;  
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; D86069; BAA13003.1;  
 DR InterPro; IPR000328; Env\_GP41.  
 DR Pfam; PF00516; GP120; 1.  
 DR Pfam; PF00517; GP41; 1.  
 KW AIDS: Coat protein; Glycoprotein; Polyprotein; Transmembrane.  
 SQ SEQUENCE 856 AA; 97396 MW; FE3E784C423C108C CRC64;

Query Match 92.4%; Score 268; DB 15; Length 856;  
 Best Local Similarity 92.3%; Pred. No. 1.7e-19;  
 Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTLEMDREINNTSLIHSLSIESQNOEKLEQELLELDKWSLMMWNI 52

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
DR AIDS: Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 856 AA; 97151 MW; C50BE038FB73659 CRC64;

Query Match
Best Local Similarity 93.1%; Score 270; DB 15; Length 856;
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLEESQNOOEKLEQELLELDKWSLMMNFI 52
DB 624 NHTTWLEMDREINNTSLIHSLEESQNOOEKLEQELLELDKWSLMMNFI 675

RESULT 9
OY 0993A6 PRELIMINARY; PRT; 645 AA.
AC 0993A6:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Truncated envelope glycoprotein (Fragment).
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1007;
RX MEDLINE=21192672; PubMed=11287644;
RA Suman S., Lockett T.D., Slobod K.S., Jones B., Riberty J.M.,
  White S.W., Doherty P.C., Hurwitz J.L.;
RT "Localization of CD4+ T cell epitope hotspots to exposed strands of
RT HIV envelope glycoprotein suggests structural influences on antigen
RT processing.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4587-4592(2001).
DR EMBL: AF321563; AAK18810.1; -
DR InterPro: IPR000328; Env-GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS: Coat protein; Envelope protein; Glycoprotein; Polypeptide;
  Transmembrane.
FT NON_TER
SQ SEQUENCE 645 AA; 72485 MW; B076514BE93362EC CRC64;

Query Match
Best Local Similarity 92.4%; Score 268; DB 15; Length 645;
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLEESQNOOEKLEQELLELDKWSLMMNFI 52
DB 593 NHTTWLEMDREINNTSLIHSLEESQNOOEKLEQELLELDKWSLMMNFI 644

RESULT 10
OY 078243 PRELIMINARY; PRT; 851 AA.
AC 078243:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Env polypeptide.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RA Carlini F., Federic M., Equestre M., Ricci S., Ratli G., Zibai Q.,
  Verani P., Rossi G.B.;
RT "Sequence analysis of HIV-1 proviral DNA from a non producer

```

```

RT chronically infected HUT-78 cellular clone.";
RN [2]
RN J. Viral Diseases 1:40-55(1992).
RP SEQUENCE FROM N.A.
RX MEDLINE=89352106; PubMed=2765297;
RA Federico M., Titti F., Butto S., Orecchia A., Carlini F., Taddeo B.,
  Macchi B., Mangiano N., Verani P., Rossi G.;
RT "Biological and molecular characterization of producer and non
RT producer clones from HUT-78 infected with a patient HIV isolate.";
RL AIDS Res. Hum. Retroviruses 5:385-396(1989).
RN [3]
RP SEQUENCE FROM N.A.
RA Titti F., Federico M., Butto S., Orecchia A., Carlini F., Taddeo B.,
  Borsetti A., Saggio I., Verani P., Rossi G.;
RT "Variability of HIV-1 virus: characteristics of an infected but not
RT productive clone.";
RL Int. J. Immunopharmacol. 3:17-23(1990).
DR EMBL: Z11530; CAA77628.1; -
DR InterPro: IPR000328; Env-GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS: Coat protein; Glycoprotein; Polypeptide; Transmembrane.
SQ SEQUENCE 851 AA; 96630 MW; 1A3767B9B7E98027 CRC64;

Query Match
Best Local Similarity 92.4%; Score 268; DB 15; Length 851;
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLEESQNOOEKLEQELLELDKWSLMMNFI 52
DB 619 NHTTWLEMDREINNTSLIHSLEESQNOOEKLEQELLELDKWSLMMNFI 670

RESULT 11
OY 085582 PRELIMINARY; PRT; 854 AA.
AC 085582:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Envelope polypeptide.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86281827; PubMed=3016298;
RA Adachi A., Gendelman H.E., Koenig S., Folks T., Willey R., Rabson A.,
  Martin M.A.;
RT "Production of acquired immunodeficiency syndrome-associated
RT retrovirus in human and nonhuman cells transfected with an infectious
RT molecular clone.";
RL J. Virol. 59:284-291(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Buckler C.E., Buckler-White A.J., Willey R.L., McCoy J.;
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buckler C.E.;
RL Submitted (JUL-1989) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=92219406; PubMed=1373204;
RA Dai L.C., Littau R., Takahashi K., Ennis F.A.;
RT "Mutation of human immunodeficiency virus type 1 at amino acid 585 on
RT gp41 results in loss of killing by CD8+ A24-restricted cytotoxic T
RT lymphocytes.";
RL J. Virol. 66:3151-3154(1992).
DR EMBL: M19921; AAA44992.1; -
DR InterPro: IPR000328; Env-GP41.

```

```

OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LM87-2;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
  Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
  "Viral variability and serum antibody response in a laboratory worker
  infected with HIV type 1 (HTLV type IIIB).";
RT AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LM87-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U12035; AAA76670.1;
DR InterPro: IPR000328; Env_GP41.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
DR AIDS: Coat protein; Glycoprotein; Polypotein; Transmembrane.
FT NON_TER
SQ SEQUENCE 752 AA; 84780 MW; 708672AD0C0E8F8 CRC64;

Query Match 95.5%; Score 277; DB 15; Length 752;
Best Local Similarity 96.2%; Pred. No. 1.7e-20;
Matches 50; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NHTTWLEMDREINNTSLHSLSIESONQOEKLELLELDKMSLWMPNI 52
Db 624 NHTTWLEMDREINNTSLHSLSIESONQOEKLELLELDKMSLWMPNI 675

RESULT 6
Q90SM7 PRELIMINARY; PRT: 856 AA.
AC Q90SM7;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HXB2;
RA Ataman-Onal Y., Cheynet V., Verrier B.;
  "Mutations and transcriptional alterations associated with the
  RT downregulation of HIV-1 envelope glycoprotein expression following
  RT acute cytopathic effects.";
RT Submited (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF358141; AAK49977.1;
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
DR AIDS: Coat protein; Glycoprotein; Polypotein; Transmembrane.
SQ SEQUENCE 856 AA; 97126 MW; 9458D02B2FD734B3 CRC64;

Query Match 95.2%; Score 276; DB 15; Length 856;
Best Local Similarity 94.2%; Pred. No. 2.5e-20;
Matches 49; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NHTTWLEMDREINNTSLHSLSIESONQOEKLELLELDKMSLWMPNI 52
Db 624 NHTTWLEMDREINNTSLHSLSIESONQOEKLELLELDKMSLWMPNI 675

RESULT 7

```

```

O89797
ID 089797 PRELIMINARY; PRT: 852 AA.
AC O89797;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LM90-2;
RX MEDLINE=95127297; PubMed=7826699;
RA Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,
  Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;
  "Viral variability and serum antibody response in a laboratory worker
  infected with HIV type 1 (HTLV type IIIB).";
RT AIDS Res. Hum. Retroviruses 10:1143-1155(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LM90-2;
RA Mulder K.E.;
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: U12035; AAA76685.1;
DR EMBL: U12036; AAA76671.1;
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
DR AIDS: Coat protein; Envelope protein; Glycoprotein; Polypotein;
  Transmembrane.
SQ SEQUENCE 852 AA; 96347 MW; 4E865229D4EB33CF CRC64;

Query Match 94.1%; Score 273; DB 15; Length 852;
Best Local Similarity 94.2%; Pred. No. 5.1e-20;
Matches 49; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 NHTTWLEMDREINNTSLHSLSIESONQOEKLELLELDKMSLWMPNI 52
Db 620 NHTTWLEMDREINNTSLHSLSIESONQOEKLELLELDKMSLWMPNI 671

RESULT 8
Q92877 PRELIMINARY; PRT: 856 AA.
AC Q92877;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Envelope glycoprotein.
GN ENV.
OS Simian-Human immunodeficiency virus.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=57667;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99098984; PubMed=9882298;
RA Cayabyab M., Karlsson G.B., Etamad-Moghadam B.A., Hofmann W.,
  Steenbeke T., Halloran M., Fanton J.W., Axheim M.K., Letvin N.L.,
  Sodroski J.G.;
  "Changes in human immunodeficiency virus type 1 envelope glycoproteins
  responsible for the pathogenicity of a multiply passaged simian-human
  immunodeficiency virus (SHIV-HXBc2).";
RT Immunodeficiency virus (SHIV-HXBc2).";
RT J. Virol. 73:976-984(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Cayabyab M., Karlsson G.B., Etamad-Moghadam B., Hofmann W.,
  Halloran M., Axheim M.W., Letvin N.L., Sodroski J.G.;
  Submited (JAN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF041850; AAD12142.1;
DR InterPro: IPR000328; Env_GP41.

```

Oy	1	NHTTLEMDREINNTSLIHSLSIESQNOEKLEDELFLDWMASLWNNFNI	52
		:     :     :     :     :     :     :     :     :     :     :	
Dd	620	NHTTWEMDRERINNTSLSIHSLSIESQNOEKNEDELLFLDKWASLWNNFNI	671
 RESULT 2			
ID	070604	PRELIMINARY; PRT: 752 AA.	
AC	070604		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
DE	Envelope glycoprotein (Fragment).		
GN	ENV.		
OS	Human immunodeficiency virus type 1.		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11676;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-LW851;		
RX	MEDLINE=95127297; PubMed=7826699;		
RA	Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B., Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.; "Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIIB)."; AIDS Res. Hum. Retroviruses 10:1143-1155(1994).		
RT	[2]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN-LW851;		
RC	Mulder K.E.;		
RL	Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: U12030; AAA76666.1; -;		
DR	InterPro: IPR000328; Env_GP41.		
DR	InterPro: IPR00777; GP120.		
DR	Pfam: PF00516; GP120; 1.		
KM	Pfam: PF00517; GP41; 1.		
KW	AIDS; Coat protein; Glycoprotein; Polyprotein; Transmembrane.		
FT	NON_TER 752		
SQ	SEQUENCE 752 AA; 84894 MW; 8B30AE894013B45A CRC64;		
 Query Match 96.9%; Score 281; DB 15; Length 752; Best Local Similarity 96.2%; Pred. No. 6,7e-21;			
Matches	50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		
Oy	1	NHTTLEMDREINNTSLSIHSLSIESQNOEKLEDELFLDWMASLWNNFNI	52
		:     :     :     :     :     :     :     :     :     :	
Dd	624	NHTTWEMDRERINNTSLSIHSLSIESQNOEKNEDELLFLDKWASLWNNFNI	675
 RESULT 3			
ID	070605	PRELIMINARY; PRT: 752 AA.	
AC	070605;		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
DE	Envelope glycoprotein (Fragment).		
GN	ENV.		
OS	Human immunodeficiency virus type 1.		
OC	Viruses; Retroid viruses; Retroviridae; Lentivirus.		
OX	NCBI_TaxID=11676;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-LW852;		
RX	MEDLINE=95127297; PubMed=7826699;		
RA	Reitz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B., Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.; "Viral variability and serum antibody response in a laboratory worker infected with HIV type 1 (HIV type IIIB)."; AIDS Res. Hum. Retroviruses 10:1143-1155(1994).		
RT	[2]		
RN	SEQUENCE FROM N.A.		
RP	STRAIN-LW852;		

Query Match	Best Local Similarity	Score	DB	Length
Matches 50; Conservative	96.2%; Pred. No. 6.7e-21;	281;	DB 15;	752;
	1; Mismatches 1; Indels 0; Gaps 0;			
<p>Query Match</p> <p>Best Local Similarity 96.9%; Score 281; DB 15; Length 752;</p> <p>Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;</p>				
<p>Query Match</p> <p>Best Local Similarity 96.2%; Pred. No. 1.7e-20;</p> <p>Matches 50; Conservative 0; Mismatches 2; Indels 0; Gaps 0;</p>				
<p>Query Match</p> <p>Best Local Similarity 95.5%; Score 277; DB 15; Length 747;</p> <p>Matches 50; Conservative 0; Mismatches 2; Indels 0; Gaps 0;</p>				

GenCore version 5.1.4.P5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:01 ; Search time 29.202 Seconds  
(without alignments)  
366.908 Million cell updates/sec

Title: US-09-877-606-9  
Perfect score: 290  
Sequence: 1 NHTTWLEWDREINNYTSLIH.....LEQELLELDKWSIAMWPNFI 52

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SPREMBL\_21:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp-organella:\*  
10: sp-phage:\*  
11: sp-plant:\*  
12: sp-rodent:\*  
13: sp-vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281	96.9	748	15 070606	070606 human immun
2	281	96.9	752	15 070604	070604 human immun
3	281	96.9	752	15 070605	070605 human immun
4	277	95.5	747	15 070607	070607 human immun
5	277	95.5	752	15 070608	070608 human immun
6	276	95.2	856	15 0905M7	0905M7 human immun
7	273	94.1	852	15 089797	089797 human immun
8	273	93.1	856	15 092877	092877 simian-huma
9	268	92.4	645	15 0933A6	0933A6 human immun
10	268	92.4	851	15 078243	078243 human immun
11	268	92.4	854	15 085582	085582 human immun
12	268	92.4	854	15 072502	072502 human immun
13	268	92.4	856	15 074599	074599 human immun
14	268	92.4	856	15 074090	074090 human immun
15	264	91.0	854	15 090178	090178 human immun
16	264	91.0	854	15 078705	078705 human immun

17	257	88.6	757	15 090722	090722 human immun
18	256	88.3	855	15 09E1R7	09E1R7 human immun
19	255	87.9	848	15 069990	069990 human immun
20	254	87.6	851	15 056110	056110 human immun
21	253	87.2	727	15 090723	090723 human immun
22	252	86.9	616	15 0993B0	0993B0 human immun
23	252	86.9	618	15 0993B2	0993B2 human immun
24	251	86.6	358	15 078120	078120 human immun
25	251	86.6	358	15 069988	069988 human immun
26	250	86.2	635	15 090U82	090U82 human immun
27	250	86.2	838	15 003806	003806 human immun
28	250	86.2	852	15 069992	069992 human immun
29	250	86.2	854	15 078225	078225 human immun
30	250	86.2	855	15 003805	003805 human immun
31	249	85.9	862	15 09E1S2	09E1S2 human immun
32	248	85.5	856	15 072993	072993 human immun
33	248	85.5	859	15 080863	080863 human immun
34	247	85.2	851	15 08Q852	08Q852 human immun
35	247	85.2	863	15 09WJ04	09WJ04 human immun
36	246	84.8	847	15 09YP48	09YP48 human immun
37	246	84.8	847	15 069996	069996 human immun
38	246	84.8	854	15 056112	056112 human immun
39	246	84.8	863	15 09WJ08	09WJ08 human immun
40	246	84.8	863	15 042031	042031 human immun
41	246	84.8	864	15 09WJ01	09WJ01 human immun
42	245	84.5	849	15 077368	077368 human immun
43	244	84.1	684	15 09IK06	09IK06 human immun
44	244	84.1	847	15 041537	041537 human immun
45	244	84.1	855	15 09IK00	09IK00 human immun

ALIGNMENTS

RESULT 1  
ID 070606 PRELIMINARY; PRT; 748 AA.  
AC 070606;  
DT 01-NOV-1996 (TRENBLREL. 01, Created)  
DT 01-NOV-1996 (TRENBLREL. 01, Last sequence update)  
DT 01-DEC-2001 (TRENBLREL. 19, Last annotation update)  
DE Envelope glycoprotein (Fragment).  
GN ENV.  
OS Human immunodeficiency virus type 1.  
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.  
OX NCBI\_TaxID=11676;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LW881;  
RX MEDLINE=95127297; PubMed=7826699;  
RA Reltz M., Hall L., Robert-Guroff M., Lautenberger J., Hahn B.,  
RA Shaw G., Kong L., Weiss S., Waters D., Gallo R., Blattner W.;  
RT "Viral variability and serum antibody response in a laboratory worker  
RT infected with HIV type 1 (HTLV type IIIB)."  
RL AIDS Res. Hum. Retroviruses 10:1143-1155(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=LW881;  
RA Mulder K.E.;  
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U12032; AAA7668.1; -;  
DR InterPro: IPR000328; Env-CP41.  
DR InterPro: IPR000777; GP120.  
DR Pfam: PF00516; GP120; 1.  
DR Pfam: PF00517; GP41; 1.  
KW AIDS; Coat protein; Glycoprotein; Polypotein; Transmembrane.  
FT NON\_TER 748  
SQ SEQUENCE 748 AA; 84224 MW; 56BEDF186C67694B CRC64;  
Query Match 96.9%; Score 281; DB 15; Length 748;  
Best Local Similarity 96.2%; Pred. No. 6.6e-21;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

**THIS PAGE BLANK (USPTO)**











[illegible]

```

FT CARBOHYD 384 384 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 390 390 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 461 461 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 609 609 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 614 614 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 623 623 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 672 672 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 748 748 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 814 814 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 853 AA; 96912 MW; 3377B993B6F22ABA CRC64;

Query Match
Best Local Similarity 91.4%; Score 265; DB 1; Length 853;
Matches 47; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLHSLSIESQNOOEKLEDELKDWASLMMWNI 52
Db 622 NNNWTWMDREINNTSLHSLSIDESQNOOEKNEDELKDWASLMMWNI 673

RESULT 8
ENV_HV1PV STANDARD; PRT; 856 AA.
ID ENV_HV1PV
AC P03376;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (PV22 isolate) (HIV-1).
OC Viruses; Retrovird viruses; Retroviridae; Lentiviruses.
OX NCBI_Taxid=11700;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8511157; PubMed=2982104;
RA Mesing M.A., Smith D.H., Cabradilla C.D., Benton C.V., Lasky L.A.,
RA Capon D.J.;
RT "Nucleic acid structure and expression of the human
Aids/Lymphadenopathy retrovirus.";
RL Nature 313:450-458(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation-
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
CC -----
DR EMBL: K02083; AAB59873.1; ALT_SEQ.
DR EMBL: X01762; CAA25903.1; ALT_SEQ.
DR PIR: A03974; VCLJVL.
DR HIV: K02083; ENVSPV22.
DR InterPro: IPR000328; Env.GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW Aids; Coat protein; Polypeptide; Glycoprotein; Transmembrane;
KW Signal.
KW SIGNAL.
FT CHAIN 1 30 EXTERIOR MEMBRANE GLYCOPROTEIN.
FT CHAIN 31 511 TRANSMEMBRANE GLYCOPROTEIN.
FT CHAIN 512 856 BY SIMILARITY.
FT DISULFID 54 74 BY SIMILARITY.
FT DISULFID 119 205 BY SIMILARITY.
FT DISULFID 126 196 BY SIMILARITY.
FT DISULFID 131 157 BY SIMILARITY.
FT DISULFID 247 247 BY SIMILARITY.
FT DISULFID 228 239 BY SIMILARITY.

```

```

FT DISULFID 296 331 BY SIMILARITY.
FT DISULFID 378 445 BY SIMILARITY.
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97339 MW; 5FCD81DC3C1209B3 CRC64;

Query Match
Best Local Similarity 89.3%; Score 259; DB 1; Length 856;
Matches 46; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLHSLSIESQNOOEKLEDELKDWASLMMWNI 52
Db 624 NNNWTWMDREINNTSLHSLSIDESQNOOEKNEDELKDWASLMMWNI 675

RESULT 9
ENV_HV1W1 STANDARD; PRT; 856 AA.
ID ENV_HV1W1
AC P31872;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (WM01 isolate) (HIV-1).
OC Viruses; Retrovird viruses; Retroviridae; Lentiviruses.
OX NCBI_Taxid=31678;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=86218077; PubMed=2423250;
RA Starcich B.R., Hahn B.H., Shaw G.M., McNeely P.D., Modrow S., Wolf H.,
RA Parks E.S., Parks W.P., Josephs S.F., Gallo R.C., Wong-Staal F.;
RT "Identification and characterization of conserved and variable
regions in the envelope gene of HTLV-III/LAV, the retrovirus of
AIDS."
FT CHAIN 45:637-648(1986).
FT CHAIN 46:637-648(1986).
CC BLOOD SAMPLES SEQUENTIALLY TAKEN FROM A TWO-YEAR OLD HAITIAN WHO
CC WAS PERINATALLY INFECTED BY HER MOTHER.
CC PIR: A24774; VCLJ3W.
DR InterPro: IPR000328; Env.GP41.
DR InterPro: IPR000777; GP120.

```



```

FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 606 606 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 620 620 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 632 632 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 669 669 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 745 745 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 811 811 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 851 AA; 96644 MW; D16A3C90857765F1 CRC64;

Query Match 92.4%; Score 268; DB 1; Length 851;
Best Local Similarity 92.3%; Pred. No. 3.7e-21;
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEMPREINNTYSLHSLIEESQNOEKLEDELLDKWASLMMNFI 52
Db 619 NNTTWEMDEINNTYSLHSLIEESQNOEKNEDELLDKWASLMMNFI 670

RESULT 5
ENV_HY1BR STANDARD; PRT; 856 AA.
AC P03375;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BH10 isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11678;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=8511123; PubMed=2578615;
RA Ratner L., Haseltine W., Patarca R., Livak K.J., Starcich B.R.,
RA Josephs S.F., Dorian E.R., Rafalski J.A., Whitehorn E.A.,
RA Baumeister K., Ivanoff L., Petteway S.R. Jr., Pearson M.L.,
RA Lautenberger J.A., Papas T.S., Chirgwin J., Chang N.T., Gallo R.C.,
RA Wong-Staal F.;
RT "Complete nucleotide sequence of the AIDS virus, HTLV-III.";
RN Nature 313:277-284(1985).
RL [2]
RN DISULFIDE BONDS, AND CARBOHYDRATE-LINKAGE SITES.
RP MEDLINE=90285159; PubMed=2355006;
RA Leonard C.K., Spellman M.W., Riddle L., Harris R.J., Thomas J.N.,
RA Gregory T.J.;
RT "Assignment of intrachain disulfide bonds and characterization of
potential glycosylation sites of the type 1 recombinant human
immunodeficiency virus envelope glycoprotein (GP120) expressed in
RT Chinese hamster ovary cells.";
RT J. Biol. Chem. 265:10373-10382(1990).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as the content is in no way
CC modified and that a statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M15654; AAA44205.1;
CC PIR: A03973; VCLJH3.
CC HIV: M15654; ENVSBH102.
CC InterPro: IPR000328; Env_GP41.

```

```

DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW AIDS; Coat protein; Polyprotein; Glycoprotein; Transmembrane;
KW signal.
FT SIGNAL 1 30
FT CHAIN 31 511
FT CHAIN 512 856
FT DISULFID 54 74
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .)
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .)
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .)
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .)
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .)
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .)
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .)
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .)
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .)
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .)
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .)
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .)
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .)
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .)
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .)
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .)
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .)
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .)
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .)
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .)
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .)
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .)
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .)
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .)
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97224 MW; 0BFEB1A18931BB27 CRC64;

Query Match 92.4%; Score 268; DB 1; Length 856;
Best Local Similarity 92.3%; Pred. No. 3.7e-21;
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEMPREINNTYSLHSLIEESQNOEKLEDELLDKWASLMMNFI 52
Db 624 NNTTWEMDEINNTYSLHSLIEESQNOEKNEDELLDKWASLMMNFI 675

RESULT 6
ENV_HY1BR STANDARD; PRT; 861 AA.
AC P03377;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope polypeptide GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
GN ENV.
OS Human immunodeficiency virus type 1 (BRU isolate) (HIV-1).
OC Viruses; Retroviral viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11686;
RN [1]

```

[illegible][illegible]



```
FT DISULFID 385 418 BY SIMILARITY.
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97212 MW; 6FAB16AF85107FE0 CRC64;

Query Match 96.9%: Score 281; DB 1; Length 856;
Best Local Similarity 96.2%: Pred. No. 1.5e-22;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWMEDREINNTSLHSLIESONQOEKLEQLELDELDMKASIMMNFNI 52
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 624 NHTTWMEDREINNTSLHSLIESONQOEKNEQLELDELDMKASIMMNFNI 675

RESULT 2
ENV_HV1H3 STANDARD: PRT: 856 AA.
AC P04624:
DT 13-AUG-1987 (Rel. 05, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
CN ENV.
OS Human immunodeficiency virus type 1 (HXB3 isolate) (HIV-1).
OC Viruses; Retrovird viruses; Retroviridae; Lentiviruses.
OX NCBI_TaxID=11707;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85228248; PubMed=2988795;
RA Crowl R., Ganguly K., Gordon M., Conroy R., Schaber M., Kramer R.,
Shaw G.M., Wong-Staal F., Reddy E.P.;
RT "HIV-1 env gene products synthesized in E. coli are recognized by
antibodies present in the sera of AIDS patients.";
RL Cell 41:979-986(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
```

```
DR EMBL: M14100; AAA44679.1; -.
DR HIV; M14100; ENVSHXB3.
DR InterPro: IPR000328; Env_GP41.
DR InterPro: IPR000777; GP120.
DR Pfam: PF00516; GP120; 1.
DR Pfam: PF00517; GP41; 1.
KW Aids; Coat protein; Glycoprotein; Polypeptide; Transmembrane;
Signal.
FT CHAIN 1 30
FT SIGNAL 31 511
FT CHAIN 512 856
FT DISULFID 119 205
FT DISULFID 126 196
FT DISULFID 131 157
FT DISULFID 218 247
FT DISULFID 228 239
FT DISULFID 296 331
FT DISULFID 378 445
FT DISULFID 385 418
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 136 136 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 141 141 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 156 156 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 160 160 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 186 186 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 197 197 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 230 230 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 234 234 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 276 276 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 289 289 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 295 295 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 386 386 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 397 397 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 448 448 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 616 616 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 624 624 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 637 637 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 674 674 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 750 750 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 816 816 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 856 AA; 97188 MW; 3373C68BB84C1AFC CRC64;

Query Match 96.9%: Score 281; DB 1; Length 856;
Best Local Similarity 96.2%: Pred. No. 1.5e-22;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWMEDREINNTSLHSLIESONQOEKLEQLELDELDMKASIMMNFNI 52
||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 624 NHTTWMEDREINNTSLHSLIESONQOEKNEQLELDELDMKASIMMNFNI 675

RESULT 3
ENV_HV1LM STANDARD: PRT: 856 AA.
AC 070626:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Envelope glycoprotein GP160 precursor [Contains: Exterior membrane
glycoprotein (GP120); Transmembrane glycoprotein (GP41)].
CN ENV.
OS Human immunodeficiency virus type 1 (LM12.3 isolate) (HIV-1).
OC Viruses; Retrovird viruses; Retroviridae; Lentiviruses.
```



RESULT 12  
T12016  
envelope glycoprotein - human immunodeficiency virus type 1 (strain sc14.3)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000  
C:Accession: T12016  
R:McCutchan, F.E.; Sanders-Buell, E.; Salminen, M.O.; Carr, J.K.; Sheppard, W.H.  
AIDS Res. Hum. Retroviruses 14, 329-337, 1998  
A:Title: Diversity of the human immunodeficiency virus type 1 envelope glycoprotein in  
A:Reference number: Z17379; MUID:9818716; PMID:9519894  
A:Accession: T12016  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-852 <MCC>  
A:Cross-references: EMBL:U09334; NID:g2351783; PIDN:AAC59271.1; PID:g2351784  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein

Query Match 81.4%; Score 236; DB 2; Length 852;  
Best Local Similarity 80.8%; Pred. No. 3.6e-17;  
Matches 42; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
QY 1 NHTTWLMDREIHNNTSLIHSLSIESQNOOEKLEQLELDKWSLMMNFI 52  
DB 620 NNMTWMEDEIHNNTSLIYTLIESQNOOEKNEQLELDKWSLMMNFI 671

RESULT 13  
S21996  
envelope protein gp120/gp41 - human immunodeficiency virus type 1 (patient 27L)  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S70422; S21996  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cerebro  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70422  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-357 <ST2>  
A:Cross-references: EMBL:X61356; NID:g60181; PIDN:CAA43624.1; PID:g1067129  
A:Experimental source: patient 27L  
A:Note: submitted to the EMBL Data Library, July 1991  
C:Superfamily: type E retrovirus env polyprotein

Query Match 81.0%; Score 235; DB 2; Length 357;  
Best Local Similarity 78.8%; Pred. No. 1.6e-17;  
Matches 41; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
QY 1 NHTTWLMDREIHNNTSLIHSLSIESQNOOEKLEQLELDKWSLMMNFI 52  
DB 125 DNMTWMEDEIHNNTSLIYTLIESQNOOEKNEQLELDKWSLMMNFI 176

RESULT 14  
S22002  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 3L  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S22002; S70418  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as determi  
A:Reference number: S21990  
A:Accession: S22002  
A:Molecule type: DNA  
A:Residues: 1-358 <STE1>  
A:Cross-references: EMBL:X61352; NID:g60186; PIDN:CAA43616.1; PID:g60187  
A:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992

A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70418  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333, 'X', 335-358 <STE2>  
A:Cross-references: EMBL:X61352; NID:g60186  
C:Superfamily: type E retrovirus env polyprotein

Query Match 81.0%; Score 235; DB 2; Length 358;  
Best Local Similarity 78.8%; Pred. No. 1.6e-17;  
Matches 41; Conservative 8; Mismatches 3; Indels 0; Gaps 0;  
QY 1 NHTTWLMDREIHNNTSLIHSLSIESQNOOEKLEQLELDKWSLMMNFI 52  
DB 126 DNMTWMEDEIHNNTSLIYTLIESQNOOEKNEQLELDKWSLMMNFI 177

RESULT 15  
S22000  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 12-Apr-1995  
C:Accession: S22000  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as dete  
A:Reference number: S21990  
A:Accession: S22000  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-358 <STE>  
A:Cross-references: EMBL:X61351  
C:Superfamily: type E retrovirus env polyprotein

Query Match 81.0%; Score 235; DB 2; Length 358;  
Best Local Similarity 78.8%; Pred. No. 1.6e-17;  
Matches 41; Conservative 8; Mismatches 3; Indels 0; Gaps 0;

QY 1 NHTTWLMDREIHNNTSLIHSLSIESQNOOEKLEQLELDKWSLMMNFI 52  
DB 126 DNMTWMEDEIHNNTSLIYTLIESQNOOEKNEQLELDKWSLMMNFI 177

Search completed: March 18, 2003, 14:37:21  
Job time : 17.6502 secs

```

R:Starich, B.R.; Hahn, B.H.; Shaw, G.M.; McNeely, P.D.; Modrow, S.; Wolf, H.; Parks, E.
Cell 45, 631-648, 1986
A:Title: Identification and characterization of conserved and variable regions in the en
A:Reference number: A24774; MUID:86218077; PMID:2423250
A:Accession: A24774
A:Molecule type: DNA
A:Residues: 1-856 <STA>
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: coat protein; glycoprotein; polyprotein; transmembrane protein
F:1-29/Domains: signal sequence #status predicted <SIG>
F:30-501/Product: coat protein gp120 #status predicted <GPI>
F:502-847/Product: coat protein gp41 #status predicted <GPI>
F:847,134,140,151,155,183,197,234,241,262,276,289,295,331,338,354,360,390,394,404,447,455
Query Match
Best Local Similarity 84.5%; Score 245; DB 1; Length 856;
Matches 43; Conservative 7; Mismatches 2; Indels 0; Gaps 0;
OY 1 NHTWLEMDREINNTSLIHSLEESONQOEKLELLELDKWSLMMFNI 52
Db 624 NNMTWMEERIDNTSLIYLLIEESONQOEKNEQELLELDKWSLMMFNI 675
RESULT 8
VCLJSC
env polyprotein precursor - human immunodeficiency virus type 1 (isolate SC)
N:Alternate names: coat polyprotein
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Feb-1997
C:Accession: B28922
R:Gungor, C.; Guo, H.G.; Franchini, G.; Aldovini, A.; Collalti, E.; Farrell, K.; Wong-Stat
Virology 164, 531-536, 1988
A:Title: Envelope sequences of two new United States HIV-1 isolates.
A:Reference number: A28922; MUID:88219542; PMID:3365091
A:Accession: B28922
A:Molecule type: DNA
A:Residues: 1-861 <GUR>
C:Genetics:
A:Gene: env
C:Superfamily: type E retrovirus env polyprotein
C:Keywords: capsid protein; coat protein; glycoprotein; polyprotein; transmembrane prote
F:1-29/Domains: signal sequence #status predicted <SIG>
F:30-861/Product: env polyprotein #status predicted <EPP>
F:87,129,135,140,143,159,163,187,198,234,241,262,276,295,301,302,333,340,356,362,386,396
Query Match
Best Local Similarity 84.1%; Score 244; DB 1; Length 861;
Matches 43; Conservative 5; Mismatches 1; Indels 0; Gaps 0;
OY 4 TWLEMDREINNTSLIHSLEESONQOEKLELLELDKWSLMMFNI 52
Db 632 TWMEERIDNTSLIYLLIEESONQOEKNEQELLELDKWSLMMFNI 680
RESULT 9
C41621
env polyprotein P - human immunodeficiency virus type 1 (fragment)
N:Alternate names: coat polyprotein
N:Contains: amino end of coat protein gp41; carboxyl end of coat protein gp120
C:Species: human immunodeficiency virus type 1, HIV-1
A:Note: host Homo sapiens (man)
C:Date: 11-Feb-1993 #sequence_revision 31-Dec-1993 #text_change 26-Aug-1999
C:Accession: C41621
R:Bugner, H.; Weiser, B.; Flaherty, K.; Guila, J.; Nguyen, P.N.; Gibbs, R.A.
Proc. Natl. Acad. Sci. U.S.A. 88, 11236-11240, 1991
A:Title: Evolution of human immunodeficiency virus type 1 nucleotide sequence diversity
A:Reference number: A41621; MUID:92107924; PMID:1765038
A:Accession: C41621
A:Molecule type: DNA

```

A:Residues: 1443 <BUS>  
 A:Cross-references: GB:M77230; NID:g3328631; PID:NAB03792.1; PID:g555015  
 A>Note: this virus was isolated from the mother's sexual partner  
 C:Genetics:  
 A:Gene: env  
 C:Superfamily: type E retrovirus env polyprotein  
 C:Keywords: AIDS: capsid protein; coat protein; glycoprotein; immunodeficiency; polyHIV  
 F:1-251/Product: coat protein gp120 (fragment) #status predicted <GP1>  
 F:252-443/Product: coat protein gp41 (fragment) #status predicted <GP2>  
 F:424-443/Domain: transmembrane #status predicted <TM>  
 F:9,23,36,48,78,101,107,131,137,143,147,153,188,200,203,351,356,365,377/Binding site

Query Match 83.1%; Score 241; DB 2; Length 443;  
 Best Local Similarity 80.8%; Pred. No. 4.8e-18;  
 Matches 42; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSIIIESQNOQKELEDELLELDKWSIWMWFI 52  
 I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
 DB 364 NNMWTWMEWREIDNTSEIYTLIESQNOQKEQELLELDKWSIWMWFSI 415

RESULT 10  
 T09448  
 envelope glycoprotein - human immunodeficiency virus type 1 (strain JRL)  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 24-Nov-1999  
 C:Accession: T09448  
 R:Yang, S.; Vinters, H.V.; Akashi, T.; O'Brien, W.A.; Chen, I.S.; Koyanagi, Y.; Namazani, A.; Cross-references: EMBL:U63632; NID:g1465777; PID:g1465781  
 A:Accession: T09448  
 A:Status: preliminary; translated from GB/EMBL/DBD  
 A:Molecule type: DNA  
 A:Residues: 1-847 <PAN>  
 A:Cross-references: EMBL:U63632; NID:g1465777; PID:g1465781  
 C:Genetics:  
 A:Gene: env  
 C:Superfamily: type E retrovirus env polyprotein

Query Match 82.1%; Score 238; DB 2; Length 847;  
 Best Local Similarity 80.8%; Pred. No. 2.2e-17;  
 Matches 42; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSIIIESQNOQKELEDELLELDKWSIWMWFI 52  
 I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
 DB 615 NNMWTWMEWREIDNTSEIYTLIESQNOQKEQELLELDKWSIWMWFDI 666

RESULT 11  
 S13289  
 - env protein - human immunodeficiency virus type 1  
 C:Species: human immunodeficiency virus type 1, HIV-1  
 C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
 C:Accession: S13289  
 R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, R.  
 Nature 348, 69-73, 1990  
 A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp120  
 A:Reference number: S13288; MUID:91043044; PMID:2172833  
 A:Accession: S13289  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-847 <OBR>  
 C:Superfamily: type E retrovirus env polyprotein

Query Match 82.1%; Score 238; DB 2; Length 847;  
 Best Local Similarity 80.8%; Pred. No. 2.2e-17;  
 Matches 42; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSIIIESQNOQKELEDELLELDKWSIWMWFI 52  
 I:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:|  
 DB 615 NNMWTWMEWREIDNTSEIYTLIESQNOQKEQELLELDKWSIWMWFDI 666

env polyprotein precursor - human immunodeficiency virus type 1 (isolate HTLV-III, BH10)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03973  
R:Ratner, L.; Haseltine, W.; Patarca, R.; Livak, K.J.; Starcich, B.; Josephs, S.F.; Doranberger, J.A.; Papas, T.S.; Ghayeb, J.; Chang, N.T.; Gallo, R.C.; Wong-Staal, F.  
Nature 313, 277-284, 1985  
A:Title: Complete nucleotide sequence of the AIDS virus, HTLV-III.  
A:Reference number: A93353; MUID:85111123; PMID:2578615  
A:Accession: A03973  
A:Molecule type: DNA  
A:Residues: 1-856 <RAT>  
A:Cross-references: GB:M15654; GB:K02008; GB:K02009; GB:K02010; NID:9326383; PIDN:AAA442  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:512-856/Product: exterior membrane glycoprotein #status predicted <TM>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,406  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status predic  
Query Match 92.4%; Score 268; DB 1; Length 856;  
Best Local Similarity 92.3%; Pred. No. 1.4e-20;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTLEMDREINNTSLIHSLEESONOQKLEQELLELDKWSLWMMFNI 52  
DB 624 NMTWMENDREINNTSLIHSLEESONOQKNEQELLELDKWSLWMMFNI 675

RESULT 4  
VCLJIV  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LAV-1a)  
N:Alternate names: coat polyprotein  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03975  
R:Wain-Hobson, S.; Sonigo, P.; Danos, O.; Cole, S.; Alizon, M.  
Cell 40, 9-17, 1985  
A:Title: Nucleotide sequence of the AIDS virus, LAV.  
A:Reference number: A90866; MUID:85099333; PMID:2981635  
A:Accession: A03975  
A:Molecule type: DNA  
A:Residues: 1-861 <WAI>  
A:Cross-references: GB:K02013; NID:9326417; PIDN:AAB59751.1; PID:9326424  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyprot  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:512-861/Product: exterior membrane glycoprotein #status predicted <TM>  
F:512-861/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,161,165,191,202,239,246,267,281,294,300,306,337,344,361,391,397,402,411  
F:616,621,630,642,679,755,821/Binding site: carbohydrate (Asn) (covalent) #status predic  
Query Match 92.4%; Score 268; DB 1; Length 861;  
Best Local Similarity 92.3%; Pred. No. 1.4e-20;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTLEMDREINNTSLIHSLEESONOQKLEQELLELDKWSLWMMFNI 52  
DB 629 NMTWMENDREINNTSLIHSLEESONOQKNEQELLELDKWSLWMMFNI 680

RESULT 5  
VCLJIV  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate LV)  
N:Alternate names: coat polyprotein

C:Species: human immunodeficiency virus type 1, HIV-1  
A:Note: host Homo sapiens (man)  
C:Date: 17-May-1985 #sequence\_revision 17-May-1985 #text\_change 16-Jul-1999  
C:Accession: A03974  
R:Ratner, L.; Smith, D.H.; Cabradilla, C.D.; Benton, C.V.; Lasky, L.A.; Capon, D.J.  
Nature 313, 450-458, 1985  
A:Title: Nucleic acid structure and expression of the human AIDS/lymphadenopathy retr  
A:Reference number: A93355; MUID:85111157; PMID:2982104  
A:Accession: A03974  
A:Molecule type: DNA  
A:Residues: 1-856 <MUS>  
A:Cross-references: GB:K02083; NID:9555008; PIDN:AAB59873.1; PID:9328559  
C:Genetics:  
A:Gene: env  
C:Superfamily: type E retrovirus env polyprotein  
C:Keywords: AIDS; capsid protein; coat protein; glycoprotein; immunodeficiency; polyp  
F:1-30/Domain: signal sequence #status predicted <SIG>  
F:512-856/Product: exterior membrane glycoprotein #status predicted <TM>  
F:512-856/Product: transmembrane glycoprotein #status predicted <TM>  
F:88,136,141,156,160,186,197,230,234,241,262,276,289,295,301,332,339,356,386,392,397,  
F:611,616,625,637,674,750,816/Binding site: carbohydrate (Asn) (covalent) #status pre  
Query Match 89.3%; Score 259; DB 1; Length 856;  
Best Local Similarity 88.5%; Pred. No. 1.3e-19;  
Matches 46; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 NHTTLEMDREINNTSLIHSLEESONOQKLEQELLELDKWSLWMMFNI 52  
DB 624 NMTWMENDREINNTSLIHSLEESONOQKNEQELLELDKWSLWMMFNI 675

RESULT 6  
S21998  
envelope protein gp120/gp41 - human immunodeficiency virus type 1  
C:Species: human immunodeficiency virus type 1, HIV-1  
A:Variety: isolate 28  
C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 01-Dec-2000  
C:Accession: S21998; S70425  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.; Hacke, W.  
submitted to the EMBL Data Library, July 1991  
A:Description: Distinct populations of HIV-1 in blood and cerebrospinal fluid as dete  
A:Reference number: S21990  
A:Accession: S21998  
A:Molecule type: DNA  
A:Residues: 1-358 <STE1>  
A:Cross-references: EMBL:X61359; NID:960182; PIDN:CAA43630.1; PID:960183  
R:Steuler, H.; Storch-Hagenlocher, B.; Wildemann, B.  
AIDS Res. Hum. Retroviruses 8, 53-59, 1992  
A:Title: Distinct populations of human immunodeficiency virus type 1 in blood and cer  
A:Reference number: S70417; MUID:92144209; PMID:1736940  
A:Accession: S70425  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-222, 'X', 224-358 <STE2>  
A:Cross-references: EMBL:X61359; NID:960182; PIDN:CAA43630.1; PID:960183  
C:Superfamily: type E retrovirus env polyprotein  
Query Match 86.6%; Score 251; DB 2; Length 358;  
Best Local Similarity 84.6%; Pred. No. 3.2e-19;  
Matches 44; Conservative 6; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTLEMDREINNTSLIHSLEESONOQKLEQELLELDKWSLWMMFNI 52  
DB 126 NMTWMENDREINNTSLIYTLIESONOQKNEQELLELDKWSLWMMFNI 177

RESULT 7  
VCLJ3W  
env polyprotein precursor - human immunodeficiency virus type 1 (isolate WMJ1)  
N:Contains: coat protein gp120; coat protein gp41  
C:Species: human immunodeficiency virus type 1, HIV-1  
C:Date: 30-Sep-1988 #sequence\_revision 30-Sep-1988 #text\_change 07-Nov-1997  
C:Accession: A24774

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:36 ; Search time 16.6502 Seconds  
(without alignments) 300.235 Million cell updates/sec

Title: US-09-877-606-9

Perfect score: 290  
Sequence: 1 NHTWLEMDREINNTSLIHSIEESONQOELELDKRWASLMMWNI 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_73:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	268	92.4	851	2 S33985	env polypeptide -
2	268	92.4	854	2 S13288	env polypeptide - huma
3	268	92.4	856	1 VCLJH3	env polypeptide pr
4	268	92.4	861	1 VCLJLV	env polypeptide pr
5	259	89.3	856	2 S21998	env polypeptide pr
6	251	86.6	358	1 VCLJ3W	env polypeptide pr
7	245	84.5	856	1 VCLJ3W	env polypeptide pr
8	244	84.1	861	1 VCLJ3C	env polypeptide pr
9	241	83.1	443	2 C41621	env polypeptide pr
10	238	82.1	847	2 T09448	env polypeptide pr
11	238	82.1	847	2 S13289	env polypeptide pr
12	236	81.4	852	2 T12016	env polypeptide pr
13	235	81.0	357	2 S21996	env polypeptide pr
14	235	81.0	358	2 S22002	env polypeptide pr
15	235	81.0	358	2 S22000	env polypeptide pr
16	233	80.3	859	1 VCLJ17	env polypeptide pr
17	233	80.3	859	1 VCLJ17	env polypeptide pr
18	232	80.0	357	2 S22006	env polypeptide pr
19	232	80.0	357	2 S21994	env polypeptide pr
20	232	80.0	853	2 S54384	env polypeptide pr
21	232	80.0	853	1 VCLJ2R	env polypeptide pr
22	231	79.7	357	2 S21992	env polypeptide pr
23	229	78.6	445	2 VCLJBR	env polypeptide pr
24	228	78.6	445	2 A41621	env polypeptide pr
25	227	78.3	729	1 S22004	env polypeptide pr
26	227	78.3	729	1 VCLJ2X	env polypeptide pr
27	227	78.3	853	1 VCLJ2J	env polypeptide pr
28	227	78.3	861	1 VCLJ2B	env polypeptide pr
29	225	77.6	846	1 VCLJND	env polypeptide pr

30	224	77.2	868	1 VCLJH4	env polypeptide -
31	223	76.9	843	1 H44001	env polypeptide pr
32	220	75.9	859	2 T01672	env polypeptide pr
33	219	75.5	136	2 J00266	env polypeptide pr
34	218	75.2	454	2 B41621	env polypeptide pr
35	214	73.8	136	2 J00934	env polypeptide pr
36	204	70.3	357	2 S21990	env polypeptide pr
37	204	70.3	856	1 A44963	env polypeptide pr
38	189	65.2	854	1 VCLJST	env polypeptide pr
39	183	63.1	877	2 S49197	env polypeptide pr
40	168	57.9	863	2 A53034	env polypeptide pr
41	122	42.1	151	2 S30450	env polypeptide -
42	121	41.7	151	2 S30448	env polypeptide -
43	121	41.7	151	2 S30452	env polypeptide -
44	121	41.7	151	2 S30451	env polypeptide -
45	121	41.7	881	1 VCLJG3	env polypeptide -

#### ALIGNMENTS

##### RESULT 1

env polypeptide - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999

C:Accession: S33985

R:Carlino, F.

submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33985

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-851 <CAR>

A:Cross-references: EMBL:211530; NID:960192; PIDN:CAA77628.1; PID:960199

C:Superfamily: type E retrovirus env polypeptide

##### Query Match

Best Local Similarity 92.4%; Score 268; DB 2; Length 851;

Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

##### RESULT 2

env polypeptide - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997

C:Accession: S13288

R:O'Brien, W.A.; Koyanagi, Y.; Namazie, A.; Zhao, J.Q.; Diagne, A.; Idler, K.; Zack, Nature 348, 69-73, 1990

A:Title: HIV-1 tropism for mononuclear phagocytes can be determined by regions of gp1

A:Reference number: S13288; MUID:91043044; PMID:2172833

A:Accession: S13288

A:Molecule type: DNA

A:Residues: 1-854 <OBR>

C:Superfamily: type E retrovirus env polypeptide

##### Query Match

Best Local Similarity 92.4%; Score 268; DB 2; Length 854;

Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

##### RESULT 3

env polypeptide - human immunodeficiency virus type 1

C:Species: human immunodeficiency virus type 1, HIV-1

C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 26-Aug-1999

C:Accession: S33985

R:Carlino, F.

submitted to the EMBL Data Library, November 1991

A:Reference number: S33979

A:Accession: S33985

A:Status: Preliminary

A:Molecule type: mRNA

A:Residues: 1-851 <CAR>

A:Cross-references: EMBL:211530; NID:960192; PIDN:CAA77628.1; PID:960199

C:Superfamily: type E retrovirus env polypeptide

##### Query Match

Best Local Similarity 92.4%; Score 268; DB 2; Length 854;

Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

DB 1 NMEDREINNYTSLHSLIEESONOQEKNEQELLELDKWSLMMWF 46

RESULT 15  
US-09-854-816-43

; Sequence 43, Application US/09854816  
; Patent No. US20020151473A1

GENERAL INFORMATION:

APPLICANT: Andrew C. Braisted

J. Kevin Judice

Robert S. McDowell

J. Christopher Phelan

Melissa A. Starovasinik

James A. Wells

TITLE OF INVENTION: Constrained Helical Peptides and Methods of

NUMBER OF SEQUENCES: 113

CORRESPONDENCE ADDRESS:

ADDRESS: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/854,816

FILING DATE: 15-May-2001

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/965,056

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Torchia, Ph.D., Timothy E.

REGISTRATION NUMBER: 36,700

REFERENCE/DOCKET NUMBER: P1005R2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 650/225-8674

TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 43:

SEQUENCE CHARACTERISTICS:

LENGTH: 269 amino acids

TYPE: Amino Acid

TOPOLOGY: Linear

SEQUENCE DESCRIPTION: SEQ ID NO: 43:

US-09-854-816-43

Query Match 84.8%; Score 246; DB 10; Length 269;

Best Local Similarity 82.7%; Pred. No. 9e-20; 3; Indels 0;

Matches 43; Conservative 6; Mismatches 3; Gaps 0;

Qy 1 NHTTWLEMDREINNYTSLHSLIEESONOQEKLEQELLELDKWSLMMWFNI 52

Db 155 NMEDREINNYTSLHSLIEESONOQEKNEQELLELDKWSLMMWFNI 206

Search completed: March 18, 2003, 14:52:58

Job time : 12.0394 secs

[illegible]

QY	5	WIEWDREINNYTSLIHSILIESQNOQEKLEQELLEDKRWASIMNMF	50
Query Match	85.2%	Score 247; DB 10; Length 46;	
Best Local Similarity	95.7%	Pred. No. 1e-20;	
Matches	44;	Conservative 1; Mismatches 1; Indels 0; Gaps 0;	



;; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS  
;; FILE REFERENCE: 22488-712  
;; CURRENT APPLICATION NUMBER: US/10/003,035  
;; CURRENT FILING DATE: 2001-11-01  
;; PRIOR APPLICATION NUMBER: 09/585,599  
;; PRIOR FILING DATE: 2000-06-02  
;; PRIOR APPLICATION NUMBER: PCT/US01/18238  
;; PRIOR FILING DATE: 2001-06-04  
;; NUMBER OF SEQ ID NOS: 75  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO: 33  
;; LENGTH: 853  
;; TYPE: PRT  
;; ORGANISM: Artificial sequence  
;; FEATURE:  
;; OTHER INFORMATION: Modified Env with multi-clade V3 loops  
US-10-003-035-33

Query Match 92.4%; Score 268; DB 9; Length 853;  
Best Local Similarity 92.3%; Pred. No. 1.2e-21;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTYSLHSLIESQNOEKLEQELLELDKWSLMMNFNI 52  
DB 773 NNTWMEMDREINNTYSLHSLIESQNOEKNEQELLELDKWSLMMNFNI 824

RESULT 9  
US-10-003-035-53  
;; Sequence 53, Application US/10003035  
;; Patent No. US2002015127A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, Danher  
;; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS  
;; FILE REFERENCE: 22488-712  
;; CURRENT APPLICATION NUMBER: US/10/003,035  
;; CURRENT FILING DATE: 2001-11-01  
;; PRIOR APPLICATION NUMBER: 09/585,599  
;; PRIOR FILING DATE: 2000-06-02  
;; PRIOR APPLICATION NUMBER: PCT/US01/18238  
;; PRIOR FILING DATE: 2001-06-04  
;; NUMBER OF SEQ ID NOS: 75  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO: 53  
;; LENGTH: 1101  
;; TYPE: PRT  
;; ORGANISM: Artificial sequence  
;; FEATURE:  
;; OTHER INFORMATION: Modified Env/Tat  
US-10-003-035-53

Query Match 92.4%; Score 268; DB 9; Length 1101;  
Best Local Similarity 92.3%; Pred. No. 1.7e-21;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTYSLHSLIESQNOEKLEQELLELDKWSLMMNFNI 52  
DB 986 NNTWMEMDREINNTYSLHSLIESQNOEKNEQELLELDKWSLMMNFNI 1037

RESULT 10  
US-10-003-035-55  
;; Sequence 55, Application US/10003035  
;; Patent No. US2002015127A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, Danher  
;; TITLE OF INVENTION: GENETIC VACCINE AGAINST HUMAN IMMUNODEFICIENCY VIRUS  
;; FILE REFERENCE: 22488-712  
;; CURRENT APPLICATION NUMBER: US/10/003,035  
;; CURRENT FILING DATE: 2001-11-01  
;; PRIOR APPLICATION NUMBER: 09/585,599  
;; PRIOR FILING DATE: 2000-06-02  
;; PRIOR APPLICATION NUMBER: PCT/US01/18238

;; PRIOR FILING DATE: 2001-06-04  
;; NUMBER OF SEQ ID NOS: 75  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO: 55  
;; LENGTH: 1186  
;; TYPE: PRT  
;; ORGANISM: Artificial sequence  
;; FEATURE:  
;; OTHER INFORMATION: Modified Env/Tat/Rev  
US-10-003-035-55

Query Match 92.4%; Score 268; DB 9; Length 1186;  
Best Local Similarity 92.3%; Pred. No. 1.8e-21;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTYSLHSLIESQNOEKLEQELLELDKWSLMMNFNI 52  
DB 986 NNTWMEMDREINNTYSLHSLIESQNOEKNEQELLELDKWSLMMNFNI 1037

RESULT 11  
US-09-854-816-19  
;; Sequence 19, Application US/09854816  
;; Patent No. US20020151473A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Andrew C. Braisted  
;; J. Kevin Judice  
;; Robert S. McDowell  
;; J. Christopher Phelan  
;; Melissa A. Starovasilk  
;; James A. Wells  
;; TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same  
;; NUMBER OF SEQUENCES: 113  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSER: Genentech, Inc.  
;; STREET: 1 DNA Way  
;; CITY: South San Francisco  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 94080  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Winpatin (Genentech)  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/09/854,816  
;; FILING DATE: 15-May-2001  
;; CLASSIFICATION: <unknown>  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/965,056  
;; FILING DATE: <unknown>  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Torchia, PhD., Timothy E.  
;; REGISTRATION NUMBER: 36,700  
;; REFERENCE/DOCKET NUMBER: P1005R2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 650/225-8674  
;; TELEFAX: 650/952-9881  
;; INFORMATION FOR SEQ ID NO: 19:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 268 amino acids  
;; TYPE: Amino Acid  
;; TOPOLOGY: Linear  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 19:  
US-09-854-816-19

Query Match 91.4%; Score 265; DB 10; Length 268;  
Best Local Similarity 90.4%; Pred. No. 7.5e-22;  
Matches 47; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTYSLHSLIESQNOEKLEQELLELDKWSLMMNFNI 52

J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 17:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 17:  
US-09-854-816-17

Query Match 92.4%; Score 268; DB 10; Length 268;  
Best Local Similarity 92.3%; Pred. No. 3.5e-22;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLSIESQNOQEKLEOELLELDKMASLMMWFI 52  
I:|||||  
Db 154 NNMTWMDREINNTSLIHSLSIESQNOQEKNEDELLELDKMASLMMWFI 205

RESULT 6  
US-09-854-816-18  
Sequence 18, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovasnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 18:  
US-09-854-816-18

Query Match 92.4%; Score 268; DB 10; Length 268;  
Best Local Similarity 92.3%; Pred. No. 3.5e-22;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLSIESQNOQEKLEOELLELDKMASLMMWFI 52  
I:|||||  
Db 154 NNMTWMDREINNTSLIHSLSIESQNOQEKNEDELLELDKMASLMMWFI 205

RESULT 7  
US-09-779-451-8  
Sequence 8, Application US/09779451  
Patent No. US20020094521A1  
GENERAL INFORMATION:  
APPLICANT: Wild, Carl T.  
TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors  
FILE REFERENCE: 1900.0300003  
CURRENT APPLICATION NUMBER: US/09/779,451  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: US 60/235,901  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/181,543  
PRIOR FILING DATE: 2000-02-10  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 8  
LENGTH: 345  
TYPE: PPT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-779-451-8

Query Match 92.4%; Score 268; DB 10; Length 345;  
Best Local Similarity 92.3%; Pred. No. 4.6e-22;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSLSIESQNOQEKLEOELLELDKMASLMMWFI 52  
I:|||||  
Db 113 NNMTWMDREINNTSLIHSLSIESQNOQEKNEDELLELDKMASLMMWFI 164

RESULT 8  
US-10-003-035-33  
Sequence 33, Application US/10003035  
Patent No. US2002015127A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Danher

LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 16:  
US-09-854-816-16

Query Match 96.9%; Score 281; DB 10; Length 268;  
Best Local Similarity 96.2%; Pred. No. 1.3e-23;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKLEQELLELDKWSLMMNFI 52  
DB 154 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 205

RESULT 2  
US-09-756-551A-8  
Sequence 8, Application US/09756551A  
Patent No. US2002051768A1  
GENERAL INFORMATION:  
APPLICANT: C. MORROW et al.  
TITLE OF INVENTION: ENCAPSULATED RECOMBINANT VIRAL  
TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 28 STATE STREET  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/756.551A  
FILING DATE: 08-JAN-2001  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/376,184  
FILING DATE: 17-AUG-1999  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/987,867  
FILING DATE: 09-DEC-1997  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/389,459  
FILING DATE: 15-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/087,009  
FILING DATE: 01-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauro, Peter C.  
REGISTRATION NUMBER: 32,360  
REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 742-4214  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: Linear  
MOLECULE TYPE: protein  
US-09-756-551A-8  
Query Match 96.9%; Score 281; DB 10; Length 519;  
Best Local Similarity 96.2%; Pred. No. 2.7e-23;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKLEQELLELDKWSLMMNFI 52  
DB 420 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 471

RESULT 3  
US-09-476-242-1  
Sequence 1, Application US/09476242  
Patent No. US20020146683A1  
GENERAL INFORMATION:  
APPLICANT: BARNETT, Susan  
APPLICANT: HARTOG, Karin  
APPLICANT: MARTIN, Eric  
TITLE OF INVENTION: MODIFIED HIV ENV POLYPEPTIDES  
FILE REFERENCE: 1605.002  
CURRENT APPLICATION NUMBER: US/09/476,242  
CURRENT FILING DATE: 1999-12-30  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 856  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-476-242-1

Query Match 96.9%; Score 281; DB 10; Length 856;  
Best Local Similarity 96.2%; Pred. No. 4.7e-23;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKLEQELLELDKWSLMMNFI 52  
DB 624 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 675

RESULT 4  
US-09-779-451-4  
Sequence 4, Application US/09779451  
Patent No. US2002094521A1  
GENERAL INFORMATION:  
APPLICANT: WILD, Carl T.  
TITLE OF INVENTION: Assay for Detection of Viral Fusion Inhibitors  
FILE REFERENCE: 1900.030003  
CURRENT APPLICATION NUMBER: US/09/779,451  
CURRENT FILING DATE: 2001-08-17  
PRIOR APPLICATION NUMBER: US 60/235,901  
PRIOR FILING DATE: 2000-09-28  
PRIOR APPLICATION NUMBER: US 60/181,543  
PRIOR FILING DATE: 2000-02-10  
NUMBER OF SEQ ID NOS: 77  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 56  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus type 1  
US-09-779-451-4

Query Match 92.4%; Score 268; DB 10; Length 56;  
Best Local Similarity 92.3%; Pred. No. 6.3e-23;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLIHSLSIESQNOEKLEQELLELDKWSLMMNFI 52  
DB 2 NHTTWLEMDREINNTSLIHSLSIESQNOEKNEQELLELDKWSLMMNFI 53

RESULT 5  
US-09-854-816-17  
Sequence 17, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted

GenCore version 5.1.4.P5\_4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:36:16 ; Search time 12.0394 Seconds  
(Without alignments)  
199.079 Million cell updates/sec

Title: US-09-877-606-9  
Perfect score: 290  
Sequence: 1 NHTTWLEMDREINNNTSLIH.....LEQELLELDKWSLMMNENI 52

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 199416 seqs, 46092074 residues  
Total number of hits satisfying chosen parameters: 199416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: Published\_Applications\_AA:\*  
2: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB pep:\*  
3: /cgn2\_6/ptodata/2/pubppaa/PCF\_NEW\_PUB pep:\*  
4: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB pep:\*  
5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB pep:\*  
6: /cgn2\_6/ptodata/2/pubppaa/US07\_PUB pep:\*  
7: /cgn2\_6/ptodata/2/pubppaa/PCFUS\_PUB pep:\*  
8: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB pep:\*  
9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUB pep:\*  
10: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB pep:\*  
11: /cgn2\_6/ptodata/2/pubppaa/US10\_PUB pep:\*  
12: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB pep:\*  
13: /cgn2\_6/ptodata/2/pubppaa/US60\_PUB pep:\*  
14: /cgn2\_6/ptodata/2/pubppaa/US60\_PUB pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281	96.9	268	10	US-09-854-816-16 Sequence 16, Appl
2	281	96.9	519	10	US-09-756-551A-8 Sequence 8, Appl
3	281	96.9	856	10	US-09-476-242-1 Sequence 1, Appl
4	268	92.4	56	10	US-09-779-451-4 Sequence 41, Appl
5	268	92.4	268	10	US-09-854-816-17 Sequence 17, Appl
6	268	92.4	268	10	US-09-854-816-18 Sequence 18, Appl
7	268	92.4	345	10	US-09-779-451-8 Sequence 33, Appl
8	268	92.4	853	9	US-10-003-035-33 Sequence 53, Appl
9	268	92.4	1101	9	US-10-003-035-53 Sequence 53, Appl
10	268	92.4	1186	9	US-10-003-035-55 Sequence 55, Appl
11	265	91.4	268	10	US-09-854-816-19 Sequence 19, Appl
12	251	86.6	269	10	US-09-854-816-46 Sequence 46, Appl
13	250	86.2	268	10	US-09-854-816-13 Sequence 13, Appl
14	247	85.2	46	10	US-09-779-451-41 Sequence 41, Appl
15	246	84.8	269	10	US-09-854-816-43 Sequence 43, Appl
16	245	84.5	268	10	US-09-854-816-9 Sequence 9, Appl
17	244	84.1	269	10	US-09-854-816-12 Sequence 12, Appl
18	243	83.8	269	10	US-09-854-816-28 Sequence 28, Appl
19	242	83.4	268	10	US-09-854-816-26 Sequence 26, Appl

20	242	83.4	619	10	US-09-891-609-4 Sequence 4, Appl
21	242	83.4	646	10	US-09-891-609-2 Sequence 2, Appl
22	242	83.4	847	10	US-09-476-242-2 Sequence 2, Appl
23	241	83.1	233	10	US-09-854-816-50 Sequence 50, Appl
24	239	82.4	269	10	US-09-854-816-44 Sequence 44, Appl
25	238	82.1	268	10	US-09-854-816-8 Sequence 8, Appl
26	237	81.7	268	10	US-09-854-816-41 Sequence 41, Appl
27	234	80.7	267	10	US-09-854-816-38 Sequence 38, Appl
28	234	80.7	269	10	US-09-854-816-32 Sequence 32, Appl
29	234	80.7	269	10	US-09-854-816-34 Sequence 34, Appl
30	234	80.7	269	10	US-09-854-816-45 Sequence 45, Appl
31	233	80.3	267	10	US-09-854-816-15 Sequence 15, Appl
32	233	80.3	269	10	US-09-854-816-6 Sequence 6, Appl
33	232	80.0	267	10	US-09-854-816-11 Sequence 11, Appl
34	232	80.0	269	10	US-09-854-816-42 Sequence 42, Appl
35	231	79.7	269	10	US-09-854-816-33 Sequence 33, Appl
36	229	79.0	268	10	US-09-854-816-10 Sequence 10, Appl
37	229	79.0	268	10	US-09-854-816-35 Sequence 35, Appl
38	229	79.0	268	10	US-09-854-816-72 Sequence 72, Appl
39	229	79.0	269	10	US-09-854-816-20 Sequence 20, Appl
40	229	79.0	269	10	US-09-854-816-31 Sequence 31, Appl
41	228	78.6	233	10	US-09-854-816-49 Sequence 49, Appl
42	227	78.3	145	12	US-10-000-321-11 Sequence 11, Appl
43	227	78.3	269	10	US-09-854-816-21 Sequence 21, Appl
44	227	78.3	269	10	US-09-854-816-25 Sequence 25, Appl
45	227	78.3	269	10	US-09-854-816-30 Sequence 30, Appl

## ALIGNMENTS

RESULT 1  
US-09-854-816-16  
Sequence 16, Application US/09854816  
Patent No. US20020151473A1  
GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
J. Kevin Judice  
Robert S. McDowell  
J. Christopher Phelan  
Melissa A. Starovastnik  
James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/854,816  
FILING DATE: 15-May-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/965,056  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, Ph.D., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:

DB 5 NNNMTWENDREINNTSLIHSLIESQNOOEKNEOELELDKWKASIMNMENTI 56

Search completed: March 18, 2003, 14:50:06  
Job time : 12.5517 secs

US-09-570-921-24  
; Sequence 24, Application US/09570921  
; Patent No. 6455265  
; GENERAL INFORMATION:  
; APPLICANT: SERRES, PIERRE-FRANCOIS  
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE  
; TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION  
; FILE REFERENCE: 106213  
; CURRENT APPLICATION NUMBER: US/09/570, 921  
; CURRENT FILING DATE: 2000-05-15  
; PRIOR APPLICATION NUMBER: PCT/FR98/02447  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: FR/97/14387  
; PRIOR FILING DATE: 1997-11-17  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 24  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Human  
US-09-570-921-24

Query Match 92.4%; Score 268; DB 4; Length 138;  
Best Local Similarity 92.3%; Pred. No. 1.3e-22;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSIESQNOQKLEOELLELDKWSLWMMFNI 52  
| : ||||||||||||||||||||||||||||||||||||||||||||  
Db 85 NNMTWMDREINNTSLIHSIESQNOQKNEOELLELDKWSLWMMFNI 136

RESULT 13  
; Sequence 26, Application US/09570921  
; Patent No. 6455265  
; GENERAL INFORMATION:  
; APPLICANT: SERRES, PIERRE-FRANCOIS  
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE  
; TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION  
; FILE REFERENCE: 106213  
; CURRENT APPLICATION NUMBER: US/09/570, 921  
; CURRENT FILING DATE: 2000-05-15  
; PRIOR APPLICATION NUMBER: PCT/FR98/02447  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: FR/97/14387  
; PRIOR FILING DATE: 1997-11-17  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 26  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Human  
US-09-570-921-26

Query Match 92.4%; Score 268; DB 4; Length 138;  
Best Local Similarity 92.3%; Pred. No. 1.3e-22;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSIESQNOQKLEOELLELDKWSLWMMFNI 52  
| : ||||||||||||||||||||||||||||||||||||||||||||  
Db 85 NNMTWMDREINNTSLIHSIESQNOQKNEOELLELDKWSLWMMFNI 136

RESULT 14  
US-09-570-921-58  
; Sequence 58, Application US/09570921  
; Patent No. 6455265  
; GENERAL INFORMATION:  
; APPLICANT: SERRES, PIERRE-FRANCOIS  
; TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE  
; TITLE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION  
; FILE REFERENCE: 106213  
; CURRENT APPLICATION NUMBER: US/09/570, 921

CURRENT FILING DATE: 2000-05-15  
; PRIOR APPLICATION NUMBER: PCT/FR98/02447  
; PRIOR FILING DATE: 1998-11-17  
; PRIOR APPLICATION NUMBER: FR/97/14387  
; PRIOR FILING DATE: 1997-11-17  
; NUMBER OF SEQ ID NOS: 144  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 58  
; LENGTH: 138  
; TYPE: PRT  
; ORGANISM: Human  
US-09-570-921-58

Query Match 92.4%; Score 268; DB 4; Length 138;  
Best Local Similarity 92.3%; Pred. No. 1.3e-22;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSIESQNOQKLEOELLELDKWSLWMMFNI 52  
| : ||||||||||||||||||||||||||||||||||||||||||||  
Db 85 NNMTWMDREINNTSLIHSIESQNOQKNEOELLELDKWSLWMMFNI 136

RESULT 15  
US-08-388-353-641  
; Sequence 641, Application US/08388353  
; Patent No. 6010895  
; GENERAL INFORMATION:

APPLICANT: Deacon, Nicholas J.  
APPLICANT: Leamont, Jennifer C.  
APPLICANT: McPhee, Dale A.  
APPLICANT: Crowe, Suzanne  
APPLICANT: Cooper, David  
TITLE OF INVENTION: NON-PATHOGENIC STRAINS OF HIV-1  
NUMBER OF SEQUENCES: 800  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Scully, Scott, Murphy & Presser  
STREET: 400 Garden City Plaza  
CITY: Garden City  
STATE: New York  
COUNTRY: United States  
ZIP: 11530

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/388, 353  
FILING DATE: 14-FEB-1995  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Digilio, Frank S.  
REGISTRATION NUMBER: 31,346  
REFERENCE/DOCKET NUMBER: 9606  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (516) 742-4343  
TELEFAX: (516) 742-4366  
TELEX: 230 901 SANS UR

INFORMATION FOR SEQ ID NO: 641:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 237 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-388-353-641

Query Match 92.4%; Score 268; DB 3; Length 237;  
Best Local Similarity 92.3%; Pred. No. 2.5e-22;  
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLIHSIESQNOQKLEOELLELDKWSLWMMFNI 52  
| : ||||||||||||||||||||||||||||||||||||||||||||

```
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 519 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-987-867A-8

Query Match
Best Local Similarity 96.9%; Score 281; DB 3; Length 519;
Pred. No. 2,2e-23;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTYSLHSLIESQNOQEKLEQELLELDKWSLMMNFI 52
DB 420 NHTTWLEMDREINNTYSLHSLIESQNOQEKNEQELLELDKWSLMMNFI 471

RESULT 9
US-07-916-098A-2
Sequence 2, Application US/07916098A
Patent No. 5871732
GENERAL INFORMATION:
APPLICANT: BURKLY, LINDA C.
APPLICANT: CHISHOLM, PATRICIA L.
APPLICANT: THOMAS, DAVID W.
APPLICANT: ROSA, MARGARET D.
APPLICANT: ROSA, JOSEPH J.
TITLE OF INVENTION: ANTI-CD4 ANTIBODY HOMOLOGS USEFUL IN
NUMBER OF SEQUENCES: 61
CORRESPONDENCE ADDRESS:
ADDRESSEE: ALLEGRETTI & WITCOFF, LTD.
STREET: 10 SOUTH WACKER DRIVE
CITY: CHICAGO
STATE: ILLINOIS
COUNTRY: U.S.A.
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: WORD PERFECT 5.1
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/916,098A
FILING DATE: July 24, 1992
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US91/08843
FILING DATE: No. 5871732ember 27, 1991
CLASSIFICATION: 424
APPLICATION NUMBER: 07/618,542
FILING DATE: No. 5871732ember 27, 1990
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: JOHN J. MC DONNELL
REGISTRATION NUMBER: 26,949
REFERENCE/DOCKET NUMBER: 92,310-G
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 715-1000
TELEFAX: (312) 715-1234
TELEX: 910/221-5317
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 856 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-916-098A-2

Query Match
Best Local Similarity 96.9%; Score 281; DB 2; Length 856;
Pred. No. 3.9e-23;
```

```
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTYSLHSLIESQNOQEKLEQELLELDKWSLMMNFI 52
DB 624 NHTTWLEMDREINNTYSLHSLIESQNOQEKNEQELLELDKWSLMMNFI 675

RESULT 10
US-09-337-387-11
Sequence 11, Application US/09337387
Patent No. 6420545
GENERAL INFORMATION:
APPLICANT: HOXIE, James A.
APPLICANT: LABRANCHE, Celia C.
APPLICANT: DOMS, Robert W.
APPLICANT: HOFFMAN, Trevor L.
TITLE OF INVENTION: CD4-INDEPENDENT HIV ENVELOPE PROTEINS AS VACCINES AND
FILE REFERENCE: HOXIE 9596-10401 (0282)
CURRENT APPLICATION NUMBER: US/09/337,387
CURRENT FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: US 09/317,556
PRIOR FILING DATE: 1999-05-24
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 11
LENGTH: 856
TYPE: PRT
ORGANISM: Human immunodeficiency virus type 1
US-09-337-387-11

Query Match
Best Local Similarity 96.9%; Score 281; DB 4; Length 856;
Pred. No. 3.9e-23;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTYSLHSLIESQNOQEKLEQELLELDKWSLMMNFI 52
DB 624 NHTTWLEMDREINNTYSLHSLIESQNOQEKNEQELLELDKWSLMMNFI 675

RESULT 11
US-09-570-921-22
Sequence 22, Application US/09570921
Patent No. 6455265
GENERAL INFORMATION:
APPLICANT: SERRES, PIERRE-FRANCOIS
TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
FILE REFERENCE: 106213
CURRENT APPLICATION NUMBER: US/09/570,921
CURRENT FILING DATE: 2000-05-15
PRIOR APPLICATION NUMBER: PCT/FR98/02447
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: FR/97/14387
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 22
LENGTH: 138
TYPE: PRT
ORGANISM: Human
US-09-570-921-22

Query Match
Best Local Similarity 92.4%; Score 268; DB 4; Length 138;
Pred. No. 1.3e-22;
Matches 48; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTYSLHSLIESQNOQEKLEQELLELDKWSLMMNFI 52
DB 85 NHTTWLEMDREINNTYSLHSLIESQNOQEKNEQELLELDKWSLMMNFI 136

RESULT 12
```

QY 1 NHTTWLEMDREINNTSLHSLIESQNOOEKLEQELLELDKWSLMMNFI 52  
 DB 420 NHTTWLEMDREINNTSLHSLIESQNOOEKLEQELLELDKWSLMMNFI 471

RESULT 6

US-08-444-882-8  
 ; Sequence 8, Application US/08444882  
 ; Patent No. 5622705  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MORROW, Casey D.  
 ; TITLE OF INVENTION: ENCAPSIDATED POLIOVIRUS NUCLEIC  
 ; TITLE OF INVENTION: ACID AND METHODS OF MAKING AND  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MASSACHUSETTS  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/444,882  
 ; FILING DATE: 19-MAY-1995  
 ; CLASSIFICATION: 514  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/087,009  
 ; FILING DATE: 01-JUL-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Geary III, William C.  
 ; REGISTRATION NUMBER: 31,359  
 ; REFERENCE/DOCKET NUMBER: UAG-004  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 519 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-444-882-8

Query Match 96.9%; Score 281; DB 1; Length 519;  
 Best Local Similarity 96.2%; Pred. No. 2.2e-23;  
 Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLHSLIESQNOOEKLEQELLELDKWSLMMNFI 52  
 DB 420 NHTTWLEMDREINNTSLHSLIESQNOOEKLEQELLELDKWSLMMNFI 471

RESULT 7

US-08-389-459A-8  
 ; Sequence 8, Application US/08389459A  
 ; Patent No. 5817512  
 ; GENERAL INFORMATION:  
 ; APPLICANT: MORROW, Casey D. and Porter, Donna C.  
 ; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT POLIOVIRUS  
 ; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
 ; NUMBER OF SEQUENCES: 23  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 60 STATE STREET, SUITE 510  
 ; CITY: BOSTON  
 ; STATE: MASSACHUSETTS

QY 1 NHTTWLEMDREINNTSLHSLIESQNOOEKLEQELLELDKWSLMMNFI 52  
 DB 420 NHTTWLEMDREINNTSLHSLIESQNOOEKLEQELLELDKWSLMMNFI 471

COUNTRY: USA  
 ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/389,459A  
 ; FILING DATE: 15-FEB-1995  
 ; CLASSIFICATION: 435  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/087,009  
 ; FILING DATE: 01-JUL-1993  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Silveri, Jean M.  
 ; REGISTRATION NUMBER: 39,030  
 ; REFERENCE/DOCKET NUMBER: UAG-004CP  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (617) 227-7400  
 ; TELEFAX: (617) 227-5941  
 ; INFORMATION FOR SEQ ID NO: 8:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 519 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-389-459A-8

Query Match 96.9%; Score 281; DB 2; Length 519;  
 Best Local Similarity 96.2%; Pred. No. 2.2e-23;  
 Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLEMDREINNTSLHSLIESQNOOEKLEQELLELDKWSLMMNFI 52  
 DB 420 NHTTWLEMDREINNTSLHSLIESQNOOEKLEQELLELDKWSLMMNFI 471

RESULT 8  
 US-08-987-867A-8  
 ; Sequence 8, Application US/08987867A  
 ; Patent No. 6063384  
 ; GENERAL INFORMATION:  
 ; APPLICANT: C. MORROW et al.  
 ; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL  
 ; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND  
 ; NUMBER OF SEQUENCES: 23  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: LAHIVE & COCKFIELD  
 ; STREET: 28 STATE STREET  
 ; CITY: BOSTON  
 ; STATE: MASSACHUSETTS  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/987,867A  
 ; FILING DATE: 09-DEC-1997  
 ; CLASSIFICATION: 424  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/087,009  
 ; FILING DATE: 01-JUL-1993  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Myers, Louis  
 ; REGISTRATION NUMBER: 35,965  
 ; REFERENCE/DOCKET NUMBER: UAG-004CPDV  
 ; TELECOMMUNICATION INFORMATION:



LENGTH: 138  
TYPE: PRT  
ORGANISM: Human  
US-09-570-921-21

Query Match  
Best Local Similarity 96.9%; Score 281; DB 4; Length 138;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLHSLIEESQNOEKLEDELLDLMKASLMMWNI 52  
DB 85 NHTTWLEMDREINNTSLHSLIEESQNOEKNEDELLDLMKASLMMWNI 136

RESULT 3  
US-08-965-056-16  
Sequence 16, Application US/08965056  
Patent No. 6271198

GENERAL INFORMATION:  
APPLICANT: Andrew C. Braisted  
APPLICANT: J. Kevin Judice  
APPLICANT: Robert S. McDowell  
APPLICANT: J. Christopher Phelan  
APPLICANT: Melissa A. Starovasnik  
APPLICANT: James A. Wells  
TITLE OF INVENTION: Constrained Helical Peptides and Methods of  
TITLE OF INVENTION: Making Same  
NUMBER OF SEQUENCES: 113  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 1 DNA Way  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Winpatlin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/965 056  
FILING DATE: 05-NO. 6271198-1997  
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:  
NAME: Torchia, PhD., Timothy E.  
REGISTRATION NUMBER: 36,700  
REFERENCE/DOCKET NUMBER: P1005R2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650/225-8674  
TELEFAX: 650/952-9881

INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-965-056-16

Query Match  
Best Local Similarity 96.9%; Score 281; DB 4; Length 268;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLHSLIEESQNOEKLEDELLDLMKASLMMWNI 52  
DB 154 NHTTWLEMDREINNTSLHSLIEESQNOEKNEDELLDLMKASLMMWNI 205

RESULT 4  
US-09-272-342B-7  
Sequence 7, Application US/09272342B  
Patent No. 6294341  
GENERAL INFORMATION:  
APPLICANT: YU, YEON-GYU

APPLICANT: KIM, SUNG-HOU  
APPLICANT: YU, JAE-REYON  
TITLE OF INVENTION: METHOD FOR DETECTING A SUBSTANCE HAVING AN ACTIVITY TO  
TITLE OF INVENTION: INHIBIT HIV INFECTION USING IMMUNOASSAY AND VARIANT  
FILE REFERENCE: 2901-0125-0  
CURRENT APPLICATION NUMBER: US/09/272,342B  
CURRENT FILING DATE: 1999-03-19  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentlin ver. 2.1  
SEQ ID NO 7  
LENGTH: 345  
TYPE: PRT  
ORGANISM: Human immunodeficiency virus  
US-09-272-342B-7

Query Match  
Best Local Similarity 96.9%; Score 281; DB 4; Length 345;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 NHTTWLEMDREINNTSLHSLIEESQNOEKLEDELLDLMKASLMMWNI 52  
DB 113 NHTTWLEMDREINNTSLHSLIEESQNOEKNEDELLDLMKASLMMWNI 164

RESULT 5  
US-08-589-446-8  
Sequence 8, Application US/08589446  
Patent No. 5614413

GENERAL INFORMATION:  
APPLICANT: Morrow, Casey D.  
TITLE OF INVENTION: ENCAPSATED POLIOVIRUS NUCLEIC  
TITLE OF INVENTION: ACID AND METHODS OF MAKING AND  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: LAHIVE & COCKFIELD  
STREET: 60 STATE STREET, SUITE 510  
CITY: BOSTON  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/589,446  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/087,009  
FILING DATE:

ATTORNEY/AGENT INFORMATION:  
NAME: Geary III, William C.  
REGISTRATION NUMBER: 31,359  
REFERENCE/DOCKET NUMBER: UAG-004  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 227-7400  
TELEFAX: (617) 227-5941  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 519 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-589-446-8

Query Match  
Best Local Similarity 96.9%; Score 281; DB 1; Length 519;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

GenCore version 5.1.4-p5.4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: March 18, 2003, 14:33:36 ; Search time 12.5317 Seconds  
(without alignments)  
121.895 Million cell updates/sec

Title: US-09-877-606-9

Perfect score: 290  
Sequence: 1 NHTTWLEMDREINNYTSLIH.....LEOELLELDKWSLMMWNI 52

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

- Issued\_Patents\_AA:\*
- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*
  - 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*
  - 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*
  - 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*
  - 5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*
  - 6: /cgn2\_6/ptodata/2/1aa/backt1test1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	281	96.9	138	4	US-09-570-921-20
2	281	96.9	138	4	US-09-570-921-21
3	281	96.9	268	4	US-08-965-056-16
4	281	96.9	345	4	US-09-272-342B-7
5	281	96.9	519	1	US-08-589-446-8
6	281	96.9	519	1	US-08-444-882-8
7	281	96.9	519	2	US-08-389-459A-8
8	281	96.9	519	2	US-08-967-867A-8
9	281	96.9	856	2	US-07-916-098A-2
10	281	96.9	856	2	US-09-337-387-11
11	268	92.4	138	4	US-09-570-921-22
12	268	92.4	138	4	US-09-570-921-24
13	268	92.4	138	4	US-09-570-921-26
14	268	92.4	138	4	US-09-570-921-58
15	268	92.4	237	3	US-08-388-353-641
16	268	92.4	237	3	US-08-488-551B-641
17	268	92.4	268	4	US-08-965-056-17
18	268	92.4	268	4	US-08-965-056-18
19	268	92.4	282	5	PCT-US95-13335-1
20	268	92.4	338	3	US-08-486-099-90
21	268	92.4	338	3	US-08-360-107A-100
22	268	92.4	338	3	US-08-484-223B-90
23	268	92.4	338	3	US-08-919-597-90
24	268	92.4	338	3	US-08-475-668A-90
25	268	92.4	338	3	US-08-465-551A-90
26	268	92.4	338	3	US-08-471-913A-90
27	268	92.4	338	4	US-08-485-264A-90

28	268	92.4	338	4	US-08-474-349A-90	Sequence 90, Appl
29	268	92.4	338	4	US-08-255-208A-26	Sequence 26, Appl
30	268	92.4	345	4	US-08-817-441-49	Sequence 49, Appl
31	268	92.4	615	4	US-09-257-490-11	Sequence 11, Appl
32	268	92.4	826	1	US-08-375-510-2	Sequence 2, Appl
33	268	92.4	826	2	US-08-487-657-2	Sequence 2, Appl
34	268	92.4	839	4	US-08-472-240A-10	Sequence 10, Appl
35	268	92.4	854	4	US-09-309-572-23	Sequence 23, Appl
36	268	92.4	856	1	US-09-124-900-9	Sequence 9, Appl
37	268	92.4	861	1	US-08-127-499A-14	Sequence 14, Appl
38	268	92.4	861	1	US-08-482-847-14	Sequence 14, Appl
39	268	92.4	861	4	US-07-956-483-10	Sequence 10, Appl
40	268	92.4	861	4	US-07-956-483-16	Sequence 16, Appl
41	268	92.4	861	4	US-08-472-240A-1	Sequence 1, Appl
42	268	92.4	861	4	US-08-472-240A-7	Sequence 7, Appl
43	268	92.4	861	4	US-08-817-441-103	Sequence 103, App
44	268	92.4	863	3	US-08-463-210-11	Sequence 11, Appl
45	268	92.4	880	2	US-08-788-815-7	Sequence 7, Appl

## ALIGNMENTS

```
RESULT 1
US-09-570-921-20
Sequence 20, Application US/09570921
Patent No. 6455265
GENERAL INFORMATION:
APPLICANT: SERRES, PIERRE-FRANCOIS
TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
FILE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
FILE REFERENCE: 106213
CURRENT APPLICATION NUMBER: US/09/570, 921
PRIOR APPLICATION NUMBER: PCT/FR98/02447
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: FR/97/14387
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 20
LENGTH: 138
TYPE: PRT
ORGANISM: Human
US-09-570-921-20
Query Match 96.9%; Score 281; DB 4; Length 138;
Best Local Similarity 96.2%; Pred. No. 5e-24;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 1 NHTTWLEMDREINNYTSLIHSLIERSONOERKLEOELLELDKWSLMMWNI 52
Db 85 NHTTWLEMDREINNYTSLIHSLIERSONOERKLEOELLELDKWSLMMWNI 136
RESULT 2
US-09-570-921-21
Sequence 21, Application US/09570921
Patent No. 6455265
GENERAL INFORMATION:
APPLICANT: SERRES, PIERRE-FRANCOIS
TITLE OF INVENTION: METHOD FOR OBTAINING VACCINES FOR PREVENTING THE
FILE OF INVENTION: PATHOGENIC EFFECTS RELATED TO A RETROVIRAL INFECTION
FILE REFERENCE: 106213
CURRENT APPLICATION NUMBER: US/09/570, 921
PRIOR APPLICATION NUMBER: PCT/FR98/02447
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: FR/97/14387
PRIOR FILING DATE: 1997-11-17
NUMBER OF SEQ ID NOS: 144
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 21
```

**THIS PAGE BLANK (USPTO)**

Search completed: March 18, 2003, 14:33:29  
Job time : 38.1429 secs

---

FT	/label= "signal_peptide"
ET	Protein
ET	31..511
ET	/label= gp120
ET	/note= "exterior membrane glycoprotein"
ET	Protein
ET	512..856
ET	/label= gp41
ET	/note= "transmembrane glycoprotein"
ET	Modified-site
ET	88
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	136
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	141
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	156
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	160
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	186
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	197
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	230
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	234
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	241
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	262
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	276
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	295
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	301
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	356
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	392
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	448
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	611
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	616
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	624
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	637
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	674
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	750
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Modified-site
ET	816
ET	/label= N-linked_glycosylation_site
ET	/note= "putative"
ET	Misc-difference
ET	289
ET	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 289"
ET	Misc-difference
ET	332
ET	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 332"
ET	Misc-difference
ET	339
ET	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 339"
ET	Misc-difference
ET	386
ET	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
ET	Misc-difference
ET	397
ET	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
ET	Misc-difference
ET	406
ET	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
ET	Misc-difference
ET	463
ET	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
ET	Misc-difference
ET	A.
ET	16-SEP-1993.
ET	XX
ET	PD
ET	XX
ET	24-FEB-1993:
ET	93MO-USO1598.
ET	PF
ET	XX
ET	PR
ET	13-MAR-1992:
ET	92US-0850770.
ET	XX
ET	(HARD ) HARVARD COLLEGE.
ET	PA
ET	Essex ME, Lee C, Lee T, Lee W;
ET	PI
ET	XX
ET	DR
ET	WPI; 1993-303140/38.
ET	XX
ET	PJ
ET	Compsn. contg. selectively de-glycosylated HIV-1 envelope protein
ET	- shows improved protective immune response
ET	PS
ET	Claim 7; Page 18, 23 and 26; 45pp; English.
ET	XX
ET	CC
ET	Mutant gp160 env protein T is specifically claimed. The
ET	amino acid sequence of the muten does not appear in the
ET	specification; sequence R41031 has been derived from the HIV-1 HXB2
ET	isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIYX,
ET	Acc.#.: P04578) and the description of T muten given in the
ET	CC specification. The combination of changes made to N-linked
ET	glycosylation sites in the C-terminal region of gp120 does not prevent
ET	viral infectivity but the resultant selective deglycosylation enables
ET	an immune response to be elicited by the muten.
ET	CC
ET	SQ
ET	Sequence 856 AA;
ET	Query Match 96.9%; Score 281; DB 14; Length 856;
ET	Best Local Similarity 96.2%; Pred.No. 1.1e-21;
ET	Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
ET	Oy
ET	1 NHITWLEMDEINNYSTLSHSLESNOOEKLEDELBDLRWASLMNFI 52
ET	
DB	624 NHITWLEMDEINNYSTLSHSLESNOOEKREBELDWRWSLMNFI 675

Query Match	Best Local Similarity	Matches	Score	DB	Length
96.9%	96.2%	50	281	14	856
Conservative	1	Mismatches	1	Indels	0
Gaps	0				0

FT	Modified-site	160	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	186	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	197	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	230	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	234	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	241	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	262	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	276	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	289	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	295	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	356	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	392	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	448	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	611	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	616	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	624	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	637	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	674	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	750	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	816	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	301	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	332	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	339	/label= "N-linked_glycosylation_site
FT		/note=	"putative"
FT	Modified-site	386	/label= "N-linked_glycosylation_site
FT		/note=	"putative"

FT	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
FT	Misc-difference	397
FT	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FT	Misc-difference	406
FT	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FT	Misc-difference	463
FT	/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
XX	W09J17705-A.	
PD	16-SEP-1993.	
XX	24-FEB-1993;	93WO-US01598.
PF	13-MAR-1992;	92US-0850770.
PR	(HARD ) HARVARD COLLEGE.	
PA	Essex ME, Lee C, Lee T, Lee W;	
PI	WPI; 1993-303140/38.	
DR	Compsn. contg. selectively de-glycosylated HIV-1 envelope protein	
XX	- shows improved protective immune response	
PT	Claim 7; Page 18, 23 and 26; 45pp; English.	
XX		
CC	Mutant gp160 env protein S is specifically claimed. The amino acid sequence of the mutcin does not appear in the CC specification; sequence R41030 has been derived from the HIV-1 HXB2 isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIV1X, Acc.#: P04578) and the description of S mutcin given in the CC specification. The combination of changes made to N-linked glycosylation sites in the C-terminal region of gp120 does not prevent viral infectivity but the resultant selective deglycosylation enables an immune response to be elicited by the mutcin.	
CC		
SQ	Sequence	856 AA;
Query Match	96.9%; Score 281; DB 14; Length 856;	
Best Local Similarity	96.2%; Pred. No. 1.le-21;	
Matches	50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
OY	1 NHTTWLEMDREINNTYSLIHSILIESQNQEKLDELDKMWASLMMWFNI 52     :	
Ddb	624 NHTTMEMDREINNTYSLIHSILIESQNQEKLDELDKMWASLMMWFNI 675	
RESULT 15		
ID	AAR41031	
XX	AAR41031 standard; protein; 856 AA.	
AC	AAR41031;	
XX	23-MAR-1994 (first entry)	
DT	Selectively deglycosylated HIV-1 HXB2 env gp160 mutcin T.	
XX	N-linked glycosylation; gp120; human immunodeficiency virus; type 1;	
KW	envelope glycoprotein; vaccine; site-directed mutagenesis.	
RW	Human immunodeficiency virus type 1 (HXB2 strain).	
XX		
OS	Key	Location/Qualifiers
PH	Peptide	1..30
FT		

FT Modified-site 262  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 276  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 289  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 295  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 301  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 356  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 392  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 448  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 611  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 616  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 624  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 637  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 674  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 750  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 816  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 332  
 FT /note= "putative"  
 FT Modified-site 339  
 FT /note= "putative N-linked glycosylation site  
 FT eliminated by substitution of Asn by  
 FT Gln at position 332"  
 FT Modified-site 386  
 FT /note= "putative N-linked glycosylation site  
 FT eliminated by substitution of Asn by  
 FT Gln at position 386"  
 FT Modified-site 397  
 FT /note= "putative N-linked glycosylation site  
 FT eliminated by substitution of Asn by  
 FT Gln at position 397"  
 FT Modified-site 406  
 FT /note= "putative N-linked glycosylation site  
 FT eliminated by substitution of Asn by  
 FT His at position 406"  
 FT Modified-site 463  
 FT /note= "putative N-linked glycosylation site  
 FT eliminated by substitution of Asn by  
 FT His at position 463"  
 FT W09317705-A.  
 FT 16-SEP-1993.  
 PD

XX 24-FEB-1993; 93WO-US01598.  
 PF  
 XX 13-MAR-1992; 92US-0850770.  
 PR  
 XX (HARD ) HARVARD COLLEGE.  
 PA  
 XX Essex ME, Lee C, Lee T, Lee W;  
 PI  
 XX WPI, 1993-303140/38.  
 DR  
 XX  
 XX Compn. contg. selectively de-glycosylated HIV-1 envelope protein  
 PT - shows improved protective immune response  
 PS  
 XX Claim 7; Page 18, 23 and 26; 45pp; English.  
 CC  
 CC Mutant gp160 env protein R is specifically claimed. The  
 CC amino acid sequence of the mutain does not appear in the  
 CC specification; sequence R41029 has been derived from the HIV-1 HXB2  
 CC isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIV1X,  
 CC Acc.#: P04578) and the description of R mutain given in the  
 CC specification. The combination of changes made to N-linked  
 CC glycosylation sites in the C-terminal region of gp120 does not prevent  
 CC viral infectivity but the resultant selective deglycosylation enables  
 CC an immune response to be elicited by the mutain.  
 CC  
 SQ Sequence 856 AA:  
 Query Match 96.9%; Score 281; DB 14; Length 856;  
 Best Local Similarity 96.2%; Pred. No. 1,le-21;  
 Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 OY 1 NHTTWMREINNTYSLTHSLIEESONOQEKLEDELTDKWSLMMNFI 52  
 DB 624 NHTTWMREINNTYSLTHSLIEESONOQEKNEDELTDKWSLMMNFI 675  
 RESULT 14  
 AAR41030  
 ID AAR41030 standard; protein; 856 AA.  
 XX  
 AC AAR41030;  
 XX  
 DF 23-MAR-1994 (first entry)  
 XX  
 DE Selectively deglycosylated HIV-1 HXB2 env gp160 mutain S.  
 XX  
 KM N-linked glycosylation; gp120; human immunodeficiency virus; type 1;  
 KM envelope glycoprotein; vaccine; site-directed mutagenesis.  
 XX  
 OS Human immunodeficiency virus type 1 (HXB2 strain).  
 XX  
 FH Key  
 FT Peptide 1..30  
 FT /label= "signal\_peptide  
 FT Protein 31..511  
 FT /label= "gp120  
 FT /note= "exterior membrane glycoprotein"  
 FT Protein 512..856  
 FT /label= "gp41  
 FT /note= "transmembrane glycoprotein"  
 FT Modified-site 88  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 136  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 141  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT Modified-site 156  
 FT /label= "N-linked\_glycosylation\_site  
 FT /note= "putative"  
 FT

FT	/label= N-linked glycosylation_site
ET	/note= "putative"
FT	Modified-site
ET	611
FT	/label= N-linked glycosylation_site
ET	/note= "putative"
FT	Modified-site
ET	616
FT	/label= N-linked glycosylation_site
ET	/note= "putative"
FT	Modified-site
ET	624
FT	/label= N-linked glycosylation_site
ET	/note= "putative"
FT	Modified-site
ET	637
FT	/label= N-linked glycosylation_site
ET	/note= "putative"
FT	Modified-site
ET	674
FT	/label= N-linked glycosylation_site
ET	/note= "putative"
FT	Modified-site
ET	750
FT	/label= N-linked glycosylation_site
ET	/note= "putative"
FT	Modified-site
ET	816
FT	/label= N-linked glycosylation_site
ET	/note= "putative"
FT	Misc-difference
ET	332
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 332"
FT	Misc-difference
ET	339
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 339"
FT	Misc-difference
ET	356
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 356"
FT	Misc-difference
ET	386
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
FT	Misc-difference
ET	397
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FT	Misc-difference
ET	406
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FT	Misc-difference
ET	463
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
FT	Misc-difference
ET	463
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
PN	W09317705-A.
XX	16-SEP-1993.
PD	24-FEB-1993; 93WO-US01598.
XX	13-MAR-1992; 92US-0850770.
PR	(HARD ) HARVARD COLLEGE.
PA	Essex ME, Lee C, Lee T, Lee W;
XX	WPI, 1993-303140/38.
DR	Compun. contg. selectively de-glycosylated HIV-1 envelope protein
XX	- shows improved protective immune response
Pt	Claim 7; Page 18, 23 and 26; 45pp: English.
XX	Mutant gp160 env protein Q is specifically claimed. The
CC	amino acid sequence of the mutlein does not appear in the

CC	specification; sequence R41028 has been derived from the HIV-1 HXB2 isolate gp160 env sequence on the SWISS-PROT database (ID = ENVS_HIV1X, Acc.#: P04578) and the description of Q mutain given in the specification. The combination of changes made to N-linked glycosylation sites in the C-terminal region of gp120 does not prevent viral infectivity but the resultant selective deglycosylation enables an immune response to be elicited by the mutain.
CC	
XX	
SQ	Sequence 856 AA:
	Query Match 96.9%; Score 281; DB 14; Length 856;
	Best Local Similarity 96.2%; Pred.No. 1,1e-21;
	Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy	1 NHHTMLEMREINNTSLIHSLEESONOKELEOELELDPKASIMNMFNI 52      :     :     :     :     :     :     :     :     :
Db	624 NHHTWMEDREINNTSLIHSLEESONOKENEDELELDPKASIMNMFNI 675
RESULT 13	
AAR41029	
ID	AAR41029 standard; protein; 856 AA.
XX	AAR41029;
AC	
XX	
DT	23-MAR-1994 (first entry)
XX	
DE	Selectively deglycosylated HIV-1 HXB2 env gp160 mutain R.
XX	
KW	N-linked glycosylation; gp120; human immunodeficiency virus; type 1;
KM	envelope glycoprotein; vaccine; site-directed mutagenesis.
XX	
OS	Human immunodeficiency virus type 1 (HXB2 strain).
XX	
FH	
Key	Location/Qualifiers
FT	Peptide 1..30
FT	/label= "signal_peptide"
FT	Protein 31..511
FT	/label= "gp120"
FT	/note= "exterior membrane glycoprotein"
FT	Protein 512..856
FT	/label= "gp41"
FT	/note= "transmembrane glycoprotein"
FT	Modified-site 88
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 136
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 141
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 156
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 160
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 186
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 197
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 230
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 234
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	Modified-site 241
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"



FT	/label= N-linked_glycosylation_site	
FT	/note= "putative"	
FT	Modified-site	
FT	750	
FT	/label= N-linked_glycosylation_site	
FT	/note= "putative"	
FT	Modified-site	
FT	816	
FT	/label= N-linked_glycosylation_site	
FT	/note= "putative"	
FT	Misc-difference	
FT	386	
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"	
FT	Misc-difference	
FT	392	
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 392"	
FT	Misc-difference	
FT	397	
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"	
FT	Misc-difference	
FT	406	
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"	
FT	Misc-difference	
FT	448	
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 448"	
FT	Misc-difference	
FT	463	
FT	/note= "putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"	
FT		
XX		
PN	W09317705-A.	
PD	16-SEP-1993.	
XX		
XX	24-FEB-1993; 93WO-US01598.	
PP		
XX	13-MAR-1992; 92US-0850770.	
PR		
XX	(HARD ) HARVARD COLLEGE.	
PA		
XX		
XX	Essex ME, Lee C, Lee T, Lee W;	
PI		
XX	WPI: 1993-303140/38.	
DR		
XX		
XX	Compsn. contg. selectively de-glycosylated HIV-1 envelope protein	
PT	- Shows improved protective Immune response	
XX		
XX	Claim 7; Page 15 and 23; 45pp; English.	
XX		
CC	Mutant gp160 env protein C6 is specifically claimed. The	
CC	amino acid sequence of the mutlein does not appear in the	
CC	specification; sequence R41027 has been derived from the HIV-1 HXB2	
CC	isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHV1X,	
CC	Acc. #: P04578) and the description of C6 mutlein given in the	
CC	specification. The combination of changes made to N-linked	
CC	glycosylation sites in the C-terminal region of gp120 does not prevent	
CC	viral infectivity but the resultant selective deglycosylation enables	
CC	an immune response to be elicited by the mutlein.	
XX		
XX	Sequence 856 AA:	
XX		

```

Query Match 96.9% Score 281 DB 14 Length 856;
Best Local Similarity 96.2% Pred. No. 1,1e-21;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 NHTTWLENDREINNTSLTSLHSLEESQNOOEKLEDELLDLKMAWLMNNT 52
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 624 NHTTWLENDREINNTSLTSLHSLEESQNOOEKLEDELLDLKMAWLMNNT 675

```

RESULT 12	
AA041028	
- ID	AA041028 standard; protein; 856 AA.
XX	
AC	AA041028;
XX	
DT	23-MAR-1994 (first entry)
XX	
DE	Selectively deglycosylated HIV-1 HXB2 env gp160 mutein Q.
XX	
KW	N-linked glycosylation; gp120; human immunodeficiency virus; type 1.
KW	envelope glycoprotein; vaccine; site-directed mutagenesis.
XX	
OS	Human immunodeficiency virus type 1 (HXB2 strain) .
XX	
FH	Key
FH	Peptide
FT	1..30
FT	/label= signal_peptide
FT	31..511
FT	/label= gp120
FT	/note= "exterior membrane glycoprotein"
FT	512..856
FT	/label= gp41
FT	/note= "transmembrane glycoprotein"
FT	88
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	136
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	141
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	156
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	160
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	186
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	197
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	230
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	234
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	241
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	262
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	276
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	289
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	295
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	301
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	392
FT	/label= N-linked_glycosylation_site
FT	/note= "putative"
FT	448
FT	Modified-site

[illegible]

FT	Protein	31.	511	/label= gp120	
FT		/note=	"exterior membrane glycoprotein"		
FT	Protein	512.	.856		
FT		/label=	gp41		
FT		/note=	"transmembrane glycoprotein"		
FT	Modified-site	88		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	136		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	141		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	156		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	160		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	186		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	197		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	230		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	234		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	241		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	262		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	276		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	289		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	295		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	301		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	332		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	339		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	356		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	611		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	616		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	624		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	637		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		
FT	Modified-site	674		/label= N-linked_glycosylation_site	
FT		/note=	"putative"		

PD		16-SEP-1993.	
XX			
PF	24-FEB-1993;	93WO-US01598.	
XX			
PR	13-MAR-1992;	92US-0850770.	
XX			
PA	(HARD ) HARVARD COLLEGE.		
XX			
PI	Essex ME, Lee C, Lee T, Lee W;		
DR	WPI; 1993-303140/38.		
XX			
PT	Compsn. contg. selectively de-glycosylated HIV-1 envelope protein		
PB	- shows improved protective immune response		
PS	Claim 7; Page 15 and 23; 45pp; English.		
XX			
CC	Mutant gp160 env protein C4 is specifically claimed. The		
CC	amino acid sequence of the mutein does not appear in the		
CC	specification; sequence R41025 has been derived from the HIV-1 HXB2		
CC	isolate gp160 env sequence on the SWISS-PROT database (ID = ENVSHIYIX,		
CC	Acc.#: P04578) and the description of C4 muten given in the		
CC	specification. The combination of changes made to N-linked		
CC	glycosylation sites in the C-terminal region of gp120 does not prevent		
CC	viral infectivity but the resultant selective deglycosylation enables		
CC	an immune response to be elicited by the mutein.		
SQ	Sequence 856 AA:		
Query Match	96.9%; Score 281; DB 14; Length 856;		
Best Local Similarity	96.2%; Pred. No. 1.1e-21;		
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
Oy	1 NHHTWMENDREINNTSLIHSLIESONOEKLEDELMLDKWASIMWNFI 52      :     :     :     :     :     :     :     :     :		
Db	624 NHHTWMENDREINNTSLIHSLIESONOEKNEDELDLKWASIMWNFI 675		
RESULT 10			
AAR41026			
ID	AAR41026 standard; protein; 856 AA.		
XX			
AC	AAR41026;		
DT	23-MAR-1994 (first entry)		
XX			
DE	Selectively deglycosylated HIV-1 HXB2 env gp160 muten C5.		
XX			
KW	N-linked glycosylation: gp120; human immunodeficiency virus; type 1;		
KW	envelope glycoprotein; vaccine; site-directed mutagenesis.		
OS	Human immunodeficiency virus type 1 (HXB2 strain).		
XX			
PH	Key	Location/Qualifiers	
FT	Peptide	1..30	
FT		/label= signal_peptide	
FT	Protein	31..511	
FT		/label= gp120	
FT		/note= "exterior membrane glycoprotein"	
FT	Protein	512..856	
FT		/label= gp41	
FT		/note= "transmembrane glycoprotein"	
FT	Modified-site	88	
FT		/label= N-linked_glycosylation_site	
FT		/note= "putative"	
FT	Modified-site	136	
FT		/label= N-linked_glycosylation_site	
FT		/note= "putative"	
FT	Modified-site	141	
FT		/label= N-linked_glycosylation_site	
FT		/note= "putative"	
FT	Modified-site	156	
FT		/label= N-linked_glycosylation_site	

[illegible]

PT	Envelope protein of acquired immune deficiency syndrome virus -
PT	useful for improved testing of human blood for antibodies against
PT	virus and as antigen for vaccines
XX	
PS	Claim 33; Fig 6A; 46pp; English.
XX	
CC	An expression vector contg. a gene (AAK60128) coding for an envelope
CC	protein of an AIDS virus, and the envelope protein of an AIDS virus
CC	(AAP60131) are claimed. The vector is pref. a member of the pEV/env
CC	family, e.g. pEV1,2 or 3/env 44-640 or 205-640.
XX	
SQ	Sequence 856 AA:
	Query Match 96.9%; Score 281; DB 7; Length 856;
	Best Local Similarity 96.2%; Pred. No. 1.le-21;
	Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY	1 NHTTWLEMDREINNTSLIHSIESNOQKELEBELLLEDKWASLWWFNI 52
	:       :       :       :       :       :
Dd	624 NHTTWMEDREINNTSLIHSLESNOQEKNEBELLEDKWASLWWFNI 675
RESULT 9	
AAR41025	ID AAR41025 standard; protein: 856 AA.
XX	AAR41025;
AC	
XX	23-MAR-1994 (first entry)
DT	
XX	Selectively deglycosylated HIV-1 HXB2 env gp160 mutein C4.
DE	
XX	N-linked glycosylation; gp120; human immunodeficiency virus; type 1;
KW	envelope glycoprotein; vaccine; site-directed mutagenesis.
KW	
XX	Human immunodeficiency virus type 1 (HXB2 strain).
OS	
XX	
FH	Key Location/Qualifiers
FH	Peptide 1..30
FT	/label= "signal_peptide"
FT	31..511
FT	/label= "gp120"
FT	/note= "exterior membrane glycoprotein"
FT	512..856
FT	/label= "gp41"
FT	/note= "transmembrane glycoprotein"
FT	88
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	136
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	141
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	156
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	160
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	186
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	197
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	230
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"
FT	234
FT	/label= "N-linked_glycosylation_site"
FT	/note= "putative"

FT	Modified-site	241	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	262	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	276	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	289	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	295	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	301	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	332	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	339	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	356	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	392	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	448	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	611	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	616	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	624	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	637	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	674	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	750	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	816	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Modified-site	386	/label= N-linked glycosylation_site
FT		/note=	"putative"
FT	Misc-difference	397	/note=
FT		/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 386"
FT	Misc-difference	406	/note=
FT		/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by Gln at position 397"
FT	Misc-difference	463	/note=
FT		/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 406"
FT	Misc-difference	463	/note=
FT		/note=	"putative N-linked glycosylation site eliminated by substitution of Asn by His at position 463"
XX	W09317705-A.		

```

AC AAR69997;
XX
XX 13-SEP-1995 (first entry)
XX
XX HIV-1 env protein.
DE
XX
XX Encapsulation; Human immunodeficiency virus; type 1; immunogenic;
XX poliovirus.
XX
XX Human immunodeficiency virus type 1.
OS
XX CA2125344-A.
XX
XX 02-JAN-1995.
PD
XX
XX 07-JUN-1994; 94CA-2125344.
PF
XX
XX 01-JUL-1993; 93US-0087009.
PR
XX (UABR-) UAB RES FOUND.
PA
XX
XX Morrow CD;
PI
XX WPI: 1995-099021/14.
DR
XX N-PSDB; AAR60575.
DR
XX
XX Method for encapsidating recombinant polio:virus nucleic acid
PT useful for providing comps. to stimulate immune response.
XX
XX
XX Disclosure; Page 38; 62pp; English.
PS
XX
XX The sequence is that of the HIV-1 env protein. The DNA encoding
CC such protein is used in a method (claimed) to encapsulate
CC poliovirus cDNA to make it more immunogenic.
CC See also AAR6995-6.
XX
XX
XX Sequence 521 AA:
SQ
Query Match 96.9%; Score 281; DB 16; Length 521;
Best Local Similarity 96.2%; Pred. No. 6.5e-22;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 NHTTWEMREINNTYSLIHSLEESONOEKLEOELLELDKWSLMMNFI 52
DB 420 NHTTWEMREINNTYSLIHSLEESONOEKNEOELLELDKWSLMMNFI 471
RESULT 7
AAM43066
ID AAM43066 standard; peptide; 853 AA.
XX
XX AAM43066;
AC
XX
XX 11-SEP-1998 (first entry)
DT
XX
XX HIV-1 gp120 protein fragment from isolate HXB2.
DE
XX
XX gp120 protein; purification; fractionation; ion exchange; chromatography;
XX binding affinity; CD4; hydrophobic interaction; size exclusion; vaccine.
XX
XX Human immunodeficiency virus type 1.
OS
XX
XX US5696238-A.
XX
XX 09-DEC-1997.
PD
XX
XX 11-MAY-1995; 95US-0439286.
PF
XX
XX 20-AUG-1991; 91US-0684963.
PR
XX 16-AUG-1993; 93US-0109002.
PR 09-MAY-1994; 94US-0240073.
PR 11-MAY-1995; 95US-0439286.
XX

```

```

PA (CHIR ) CHIRON CORP.
XX
XX Haigwood NL, Scandella C;
PI
XX
XX WPI: 1998-041353/04.
DR
XX
XX Purification of HIV gp120 - using chromatographic methods
PT
XX
XX Disclosure; Fig 2A-W; 53pp; English.
PS
XX
XX AAM43066-W43080 are fragments of the gp120 protein from different human
CC immunodeficiency virus type I (HIV-1) isolates. These proteins are used
CC in a novel method for purifying HIV gp120 so as to provide a purified
CC gp120 glycoprotein having protein/protein binding properties
CC substantially identical to natural viral HIV gp120. The method involves
CC fractionating a crude gp120 preparation containing full-length,
CC glycosylated gp120 using ion exchange chromatography so as to provide a
CC first collection of fractions. A fraction from the first collection is
CC selected that exhibits specific binding affinity for CD4 peptide,
CC thereby producing a first fractionated material. The first fractionated
CC material is fractionated by hydrophobic interaction chromatography so as
CC to provide a second collection of fractions from which a second
CC collection is selected that exhibits specific binding affinity for CD4
CC peptide. This second fraction is fractionated by size exclusion
CC chromatography so as to provide a third collection of fractions
CC exhibiting specific binding affinity for CD4 peptide, thereby providing
CC the purified gp120. The purified gp120 can be used for antibody
XX production and in vaccines.
XX
XX
XX Sequence 853 AA:
SQ
Query Match 96.9%; Score 281; DB 19; Length 853;
Best Local Similarity 96.2%; Pred. No. 1.1e-21;
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
OY 1 NHTTWEMREINNTYSLIHSLEESONOEKLEOELLELDKWSLMMNFI 52
DB 621 NHTTWEMREINNTYSLIHSLEESONOEKNEOELLELDKWSLMMNFI 672
RESULT 8
AAP60131
ID AAP60131 standard; Protein; 856 AA.
XX
XX AAP60131;
AC
XX
XX 26-JUN-1991 (first entry)
DT
XX
XX Sequence of the AIDS envelope protein.
DE
XX
XX AIDS; HIV; LAV; HTLV-III; vaccine; antibody; epitope; antigen;
XX diagnosis.
XX
XX HTLV-III.
OS
XX
XX EP199301-A.
XX
XX 29-OCT-1986.
PD
XX
XX 18-APR-1986; 86EP-0105371.
PF
XX
XX 19-APR-1985; 85US-0725021.
PR
XX
XX (HOFF ) HOFFMANN-LA ROCHE AG.
XX (USGO ) US GOVERNMENT.
XX (HEAL-) DEPT. HEALTH & HUMAN SERV.
XX (USDR ) US DEPT HEALTH & HUMAN.
XX
XX Crowl RM, Gallo RC, Reddy EP, Shaw GM, Wong-Staal FY;
PI WPI: 1986-286067/44.
DR N-PSDB; AAP60128.
XX

```

PT vaccines against human immune deficiency virus, mimics the intermediate  
PT state of the native polypeptide -  
XX  
PS Disclosure: Page 29; 29pp; French.  
XX  
CC The present invention relates to novel mutant HIV gp41 LAI proteins  
CC (AB883411-AB883420). The mutants can form a structure corresponding to,  
CC or mimicking, the intermediate state of gp41. The mutants, or their  
CC conjugate with a carrier, or vectors containing nucleic acid that encode  
CC them, are used in vaccines for treatment or prevention of infection by  
CC HIV. The present sequence is wild-type HIV gp41 LAI protein, which  
CC was used to generate the mutants of the invention.  
XX  
SQ Sequence 344 AA:  
  
Query Match 96.9%; Score 281; DB 23; Length 344;  
Best Local Similarity 96.2%; Pred. No. 4,1e-22;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 NHTTWLEMDREINNTSLIHSLSIESONQOEKLEDELLDLKWSLMMNFNI 52  
IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:  
DB 113 NHTTWLEMDREINNTSLIHSLSIESONQOEKNEDELLDLKWSLMMNFNI 164  
  
RESULT 4  
AAR53785  
ID AAR53785 standard; Protein: 420 AA.  
XX  
AC AAR53785;  
XX  
DT 28-DEC-1994 (first entry)  
XX  
DE Translation of HIV-1 in the region encoding the gp41 polypeptide.  
XX  
KM Human immunodeficiency virus; HIV-1; AIDS; lentivirus; TM protein;  
KM extracellular protein; transmembrane protein; gp41.  
XX  
OS HIV-1.  
XX  
FH Key Location/Qualifiers  
FT Protein 43..387  
FT /label= gp41  
FT  
XX  
PN WO9412533-A.  
XX  
PD 09-JUN-1994.  
XX  
PF 12-JAN-1993; 93WO-US00212.  
XX  
PR 23-NOV-1992; 92US-0979975.  
XX  
PA (HARD ) HARVARD COLLEGE.  
XX  
PI Essex ME, Lee TH, Yu X;  
PI  
DR WPI; 1994-200197/24.  
DR N-PSDB; AA066275.  
XX  
PT Method for treating HIV patients - comprises administration of  
PT mutated gp41 polypeptide  
XX  
PS Claim 11; Fig 1; 54pp; English.  
XX  
CC The inventors claim a method of treating a patient infected with HIV  
CC by administering a mutated gp41 polypeptide or a therapeutic  
CC composition comprising nucleic acid encoding the mutant gp41  
CC polypeptide in an expressible genetic construction. The mutant gp41  
CC polypeptide contains a deletion of at least one AA in at least one  
CC of the following regions of wild type gp41 (AAR53783): AAs 844-856;  
CC 814-856; 796-856; 776-856; 753-856; or 710-856, effective to either  
CC disrupt viral replication or HIV or disrupt the assembly of viral  
CC Env proteins in an HIV infected cell. AA066275 corresp. to bps 7631-  
CC 8890 of wt HIV-1. X in the AA sequence represents the posn. of a

CC stop codon in AA066275.  
XX  
SQ Sequence 420 AA:  
  
Query Match 96.9%; Score 281; DB 15; Length 420;  
Best Local Similarity 96.2%; Pred. No. 5,1e-22;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 NHTTWLEMDREINNTSLIHSLSIESONQOEKLEDELLDLKWSLMMNFNI 52  
IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:  
DB 155 NHTTWLEMDREINNTSLIHSLSIESONQOEKNEDELLDLKWSLMMNFNI 206  
  
RESULT 5  
AAW00181  
ID AAW00181 standard; Protein: 519 AA.  
XX  
AC AAW00181;  
XX  
DT 12-NOV-1996 (first entry)  
XX  
DE HIV-1 env protein.  
XX  
KM Poliovirus; encapsidation; capsid; vaccine; genetic immunisation;  
KM HIV-1; human immunodeficiency virus type 1; env protein.  
XX  
OS Human immunodeficiency virus type 1.  
XX  
PN WO9625173-A1.  
XX  
PD 22-AUG-1996.  
XX  
PF 13-FEB-1996; 96WO-US01895.  
XX  
PR 15-FEB-1995; 95US-0389459.  
XX  
PA (UABR-) UAB RES FOUND.  
XX  
PI Ansardi DC, Morrow CD, Porter DC;  
PI  
DR WPI; 1996-393136/39.  
DR N-PSDB; AAT33295.  
XX  
PT Encapsidation of recombinant polio:virus nucleic acid for use in  
PT vaccines - using a polio:virus nucleic acid which lacks the P1  
PT capsid region and an expression system which provides the region  
XX  
PS Disclosure: Page 61-63; 108pp; English.  
XX  
CC cDNA sequences (AAT33293-95) respectively code for the gag, pol and  
CC env proteins (AAW00179-81) of HIV-1. They can be used to substitute  
CC the P1 capsid gene of poliovirus in recombinant poliovirus nucleic  
CC acids (rPNAs). Such rPNAs are encapsidated by introduction into  
CC a host cell together with a vaccinia virus or plasmid vector  
CC encoding the poliovirus P1 capsid precursor protein. Encapsidated  
CC rPNAs are useful for genetic immunisation, stimulating an immune  
CC response to the HIV-1 protein.  
XX  
SQ Sequence 519 AA:  
  
Query Match 96.9%; Score 281; DB 17; Length 519;  
Best Local Similarity 96.2%; Pred. No. 6,5e-22;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 NHTTWLEMDREINNTSLIHSLSIESONQOEKLEDELLDLKWSLMMNFNI 52  
IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:IIIIII:  
DB 420 NHTTWLEMDREINNTSLIHSLSIESONQOEKNEDELLDLKWSLMMNFNI 471  
  
RESULT 6  
AAR69997  
ID AAR69997 standard; Protein: 521 AA.  
XX

PT Production of constrained helical peptide(s) by linking side chains  
PT on termini of octa-peptide - derived from human immunodeficiency  
PT virus gp41 protein, useful in vaccines for treatment and prevention  
PT of infection  
XX  
PS Claim 11; Page 156; 279pp; English.  
XX  
CC Peptides AAY22805-Y22917 are derived from Human immunodeficiency virus  
CC (HIV). Specifically, AAY22810-Y22910 are derived from gp41 proteins  
CC of known HIV virus strains (AAY22871, AAY22880, AAY22888 and  
CC AAY22903 represent consensus sequences of various sections of the gp41  
CC protein). Sequences derived from the peptides are used to produce  
CC constrained helical peptides of the invention. The constrained helical  
CC peptide is produced by synthesizing an octapeptide in which both terminal  
CC amino acids have a side-chain that includes a group able to form an amide  
CC bond, and cyclizing the octapeptide by reacting the specified side-chain  
CC residues with a difunctional linker to produce two amide bonds.  
CC The constrained helical peptides are used to treat or prevent HIV  
CC infection, especially as vaccines that generate antibodies that  
CC prevent viral membrane fusion or infectivity. Vaccines may contain  
CC constrained helical peptides derived from several different strains of  
CC HIV. The antibodies are also useful for diagnosing HIV infection. Other  
CC uses for the constrained helical peptides are in affinity purification  
CC of ligands (particularly where complete binding protein is not readily  
CC available, e.g. replacements for protein A in immunoglobulin  
CC purification); as epitope mimics for antibody production; for isolation  
CC of synthetic antibody clones from phage display libraries, or as stable  
CC forms of "floppy" peptides or proteins.  
SQ Sequence 268 AA:  
  
Query Match 96.9%; Score 281; DB 19; Length 268;  
Best Local Similarity 96.2%; Pred. No. 3.1e-22;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
OY 1 NHTTWLEMDREINNTSLIHSIESQNOEKLEQELLELDKWSIMWNFI 52  
DB 154 NHTTWLEMDREINNTSLIHSIESQNOEKNEQELLELDKWSIMWNFI 205  
  
RESULT 2  
ABG68291  
ID ABG68291 standard; Protein; 268 AA.  
XX  
AC ABG68291;  
XX  
DT 07-OCT-2002 (first entry)  
XX  
DE Envelope protein gp41 from HIV clade B strain #10.  
XX  
KM HIV, glycoprotein: gp41; antigen; helical conformation;  
KM virus-induced membrane fusion; acquired immunodeficiency syndrome; AIDS;  
KM viral envelope protein; vaccine; virucide; anti-HIV.  
XX  
OS Human immunodeficiency virus type 1 clade B.  
XX  
PN US6271198-B1.  
XX  
PD 07-AUG-2001.  
XX  
PF 05-NOV-1997; 97US-0965056.  
XX  
PR 16-JUN-1997; 97US-049787P.  
PR 06-NOV-1996; 96US-0743698.  
PR 16-JUN-1997; 97US-0876698.  
XX  
PA (GENENTECH INC.  
XX  
PI Bralsted AC, Judice JK, McDowell RS, Pheelan JC, Starovasnik MA;  
PI Wells JA;  
XX  
DR WPI: 2002-487624/52.  
XX

PT New cyclic peptides from human immune deficiency virus gp41, useful for  
PT treatment or prevention of HIV infection, are constrained to have  
PT alpha-helical conformation -  
XX  
PS Disclosure: Column 151-154; 175pp; English.  
XX  
XX The invention relates to cyclic peptides (A) with a constrained helical  
CC conformation, derived from gp41 (glycoprotein 41, a viral envelope  
CC protein) of human immunodeficiency virus (HIV). The cyclic  
CC peptides have formulas given in the specification part of which are  
CC derived from a consensus sequence of gp41 derived from HIV clades A, B,  
CC C, D, E or O. The peptides are used to cause induction of a specific  
CC immune response, resulting in antibodies that prevent virus-induced  
CC membrane fusion. The peptides are used to treat subjects with, or at risk  
CC of, HIV infection, either as antifusion/anti-infection agents or,  
CC preferably where associated with a carrier, as an immunogen (including as  
CC vaccine) to raise antibodies. The antibodies may be used for diagnosis or  
CC prevention/treatment of HIV infection (i.e. acquired immunodeficiency  
CC syndrome, AIDS), e.g. prevention of mother-to-child transmission or in  
CC cases of health care accidents. The peptides can be based on specific HIV  
CC strains, e.g. breakthrough isolates of HIV that have developed during  
CC vaccine trials, so a combination of them should cover a wide range of  
CC protection. The present sequence is gp41 protein from a particular  
CC HIV clade used to derive a consensus sequence of gp41.  
SQ Sequence 268 AA:  
  
Query Match 96.9%; Score 281; DB 23; Length 268;  
Best Local Similarity 96.2%; Pred. No. 3.1e-22;  
Matches 50; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
OY 1 NHTTWLEMDREINNTSLIHSIESQNOEKLEQELLELDKWSIMWNFI 52  
DB 154 NHTTWLEMDREINNTSLIHSIESQNOEKNEQELLELDKWSIMWNFI 205  
  
RESULT 3  
ABB83400  
ID ABB83400 standard; protein; 344 AA.  
XX  
AC ABB83400;  
XX  
DT 19-SEP-2002 (first entry)  
XX  
DE HIV gp41 LAI protein #1.  
XX  
KM HIV, gp41; LAI; virucide; anti-HIV; vaccine; HIV infection.  
XX  
OS Human immunodeficiency virus.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..23  
FT /label= Fusion\_peptide 173..194  
FT Domain /label= Transmembrane\_domain  
XX  
XX WO200253587-A2.  
XX  
PN 11-JUL-2002.  
XX  
PD 04-JAN-2002; 2002WO-FR00031.  
XX  
PF 05-JAN-2001; 2001FR-0000141.  
XX  
PR 23-JAN-2001; 2001FR-0000848.  
XX  
PA (AVERT ) AVENTIS PASTEUR.  
XX  
PI Brasseur R, Charlotaux B, Chevallier M, El Habib R, Krell T;  
PI Sodoyer R;  
XX  
DR WPI: 2002-528852/56.  
XX  
PT New mutant human immune deficiency virus gp41 polypeptide, useful in

GenCore version 5.1.4.p5.4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:27:06 ; Search time 37.1429 Seconds  
(without alignments)  
186.551 Million cell updates/sec

Title: US-09-877-606-9

Perfect score: 290

Sequence: 1 NHHTTWLEMDREINNYSLIH.....LEQELLELDKWSLMMWENI 52

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_101002.\*  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	281	96.9	268	19	AA122820	SEQ ID NO. 16 from
2	281	96.9	268	23	ABG68291	Envelope protein g
3	281	96.9	344	23	ABB83400	HIV gp41 LAI prote
4	281	96.9	420	15	AA53785	Translation of HIV
5	281	96.9	519	17	AA00181	HIV-1 env protein.
6	281	96.9	521	16	AA69997	HIV-1 env protein.
7	281	96.9	833	19	AA43066	HIV-1 gp120 protei
8	281	96.9	856	7	AA60131	Sequence of the AI
9	281	96.9	856	14	AA41025	Selectively deglyc
10	281	96.9	856	14	AA41026	Selectively deglyc

11	281	96.9	856	14	AA41027	Selectively deglyc
12	281	96.9	856	14	AA41028	Selectively deglyc
13	281	96.9	856	14	AA41029	Selectively deglyc
14	281	96.9	856	14	AA41030	Selectively deglyc
15	281	96.9	856	14	AA41031	Selectively deglyc
16	281	96.9	856	14	AA41032	Selectively deglyc
17	281	96.9	856	21	AA97072	Wld type HIV-1 HX
18	281	96.9	856	22	AA845697	HIV-1/IIIB env clo
19	268	92.4	56	21	AA14532	HIV-1 isolate LAI
20	268	92.4	56	22	AA070191	HIV viral envelope
21	268	92.4	56	22	AA63860	Amino acid sequenc
22	268	92.4	138	21	AA969792	HIV-1 gp41 envelop
23	268	92.4	150	9	AA80745	Sequence of Aas 60
24	268	92.4	177	23	ABB83401	HIV gp41 LAI prote
25	268	92.4	177	23	ABB83413	HIV gp41 LAI prote
26	268	92.4	192	20	AA124098	Recombinant gp41 p
27	268	92.4	233	17	AA93199	Recombinant HIV pr
28	268	92.4	268	19	AA122821	SEQ ID NO. 17 from
29	268	92.4	268	19	AA122822	SEQ ID NO. 18 from
30	268	92.4	268	23	ABG68292	Envelope protein g
31	268	92.4	268	23	ABG68293	Envelope protein g
32	268	92.4	342	13	AA826787	ENV9. Synthetic.
33	268	92.4	344	23	ABB83410	HIV gp41 LAI prote
34	268	92.4	345	21	AA814536	HIV-1 isolate LAI
35	268	92.4	345	22	AA63863	Amino acid sequenc
36	268	92.4	393	15	AA825687	gp41 HIV envelope
37	268	92.4	393	15	AA845445	Sequence encoded b
38	268	92.4	395	10	AA91956	HIV p41 gene with
39	268	92.4	709	10	AA95781	Fusion protein of
40	268	92.4	751	10	AA90604	HIV-1 env segment
41	268	92.4	851	9	AA80967	HIV protein HT6.
42	268	92.4	854	21	AA810697	HIV-1 env protein.
43	268	92.4	854	21	AA810053	HIV-1 coat protein
44	268	92.4	854	22	AA861399	HIV gp41 DNA SEQ I
45	268	92.4	854	22	AA867277	Protein encoded by

# ALIGNMENTS

RESULT 1	
AA122820	
ID	AA122820 standard; Protein: 268 AA.
XX	
AC	AA122820;
XX	
DT	19-AUG-1999 (first entry)
XX	
DE	SEQ ID NO. 16 from WO9820036.
XX	
KW	HIV: gp41 protein; constrained helical peptide; HIV infection;
KW	vaccine; antibody; viral membrane fusion; viral infectivity;
KW	ligand affinity purification; protein A replacement;
KW	immunoglobulin purification; epitope mimc.
XX	
OS	Human immunodeficiency virus.
XX	
PN	WO9820036-A1.
XX	
PD	14-MAY-1998.
XX	
PE	05-NOV-1997; 97WO-US20069.
XX	
PR	16-JUN-1997; 97US-0876698.
XX	
PR	06-NOV-1996; 96US-0743698.
XX	
PA	(GETH ) GENENTECH INC.
XX	
PI	Braisted A, Judice JK, McDowell RS, Phelan JC, Starovasnik MA;
PI	Wells JA;
DR	WPI; 1998-286866/25.



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 25, 2003, 12:52:44 ; Search time 197.31 Seconds  
(without alignments)  
574.569 Million cell updates/sec

Title: US-09-877-606-8  
Perfect score: 7  
Sequence: 1 xxxxxx 7

Scoring table: BIOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xih  
-O/cgn2\_1/USPTO.spool/US09877606/runat.18032003.134800.17705/app\_query.fasta.1.1059  
-DB=EST -QPM=fastap -SUPPLY=1st -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTPM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USFR=US09877606.@CGN.1.1.4134.@runat.18032003.134800.17705 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGEOUTR -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estlm:\*  
5: em\_estlov:\*  
6: em\_estlpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vtl:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	0	0.0	1	2	HS0001740
2	0	0.0	1	2	HS0001740
3	0	0.0	1	2	HS0003869
4	0	0.0	1	2	HS0003869
5	0	0.0	1	2	HS0005008
6	0	0.0	1	2	HS0005008
7	0	0.0	1	2	HS0011115
8	0	0.0	1	2	HS0011115
9	0	0.0	1	2	HS0011270
10	0	0.0	1	2	HS0011270
11	0	0.0	1	2	HS0011922
12	0	0.0	1	2	HS0011922
13	0	0.0	1	13	BI416880
14	0	0.0	1	13	BI416880
15	0	0.0	1	13	BI817896
16	0	0.0	1	13	BI817896
17	0	0.0	1	14	BM817067
18	0	0.0	1	14	BM817067
19	0	0.0	1	17	AZ363337
20	0	0.0	1	17	AZ363337
21	0	0.0	2	2	HS0003817
22	0	0.0	2	2	HS0003817
23	0	0.0	2	2	HS0003931
24	0	0.0	2	2	HS0003931
25	0	0.0	2	2	HS0007187
26	0	0.0	2	2	HS0007187
27	0	0.0	2	2	HS0008709
28	0	0.0	2	2	HS0008709
29	0	0.0	2	2	HS0011919
30	0	0.0	2	2	HS0011919
31	0	0.0	2	13	BI817789
32	0	0.0	2	13	BI817789
33	0	0.0	2	14	BM817265
34	0	0.0	2	14	BM817265
35	0	0.0	2	14	C55081
36	0	0.0	2	14	C55081
37	0	0.0	2	17	AZ463604
38	0	0.0	2	17	AZ463604
39	0	0.0	2	17	BH754642
40	0	0.0	2	17	BH754642
41	0	0.0	3	2	HS0003852
42	0	0.0	3	2	HS0003852
43	0	0.0	3	2	HS0007999
44	0	0.0	3	2	HS0007999
45	0	0.0	3	2	HS0008070

## ALIGNMENTS

**THIS PAGE BLANK (USPTO)**

RL J. Bacteriol. 177:7131-7140(1995).  
 DR EMBL; X57583; CAA40808.1; -  
 KW Plasmid.  
 SQ SEQUENCE 7 AA; 763 MW; 644DD44861B406F0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
 DB 3 T 3

## RESULT 14

P70804 PRELIMINARY; PRT; 7 AA.  
 AC P70804;  
 DT 01-FEB-1997 (TREMBLrel. 02, Created)  
 DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Algt protein (Fragment).  
 GN ALGT.  
 OS Azotobacter vinelandii.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 CC Azotobacter.  
 OX NCBI\_TaxID=354;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=E;  
 RX MEDLINE=96427318; PubMed=8830682;  
 RA Rehm B.H.A., Ertesvag H., Valla S.;  
 RT "A new Azotobacter vinelandii mannuronan C-5-epimerase gene (algt) is  
 to that in Pseudomonas aeruginosa."  
 RL J. Bacteriol. 178:5884-5889(1996).  
 DR EMBL; X87973; CAA61230.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 684 MW; 71B5A5A5A2D1AED0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
 DB 2 T 2

## RESULT 15

O54248 PRELIMINARY; PRT; 7 AA.  
 AC O54248;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE RplO protein (Fragment).  
 GN RplO.  
 OS Streptomyces griseus.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 CC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1911;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N2-3-11;  
 RX MEDLINE=20011291; PubMed=10542330;  
 RA Poehling S., Piepersberg W., Wehmeier U.F.;  
 RT "Analysis and regulation of the sec Y gene from Streptomyces griseus  
 N2-3-11 and interaction of the SecY protein with the SecA protein.";  
 RL Biochim. Biophys. Acta 1447:298-302(1999).  
 DR EMBL; X95915; CAA65160.1; -  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 760 MW; 72C72B01B2D1B2A0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
 DB 2 T 2

Search completed: March 18, 2003, 14:36:08  
 Job time : 5.93103 secs

RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94079349; PubMed=8257126;  
 RA Rather P.N., Mann P.A., Miezwa R., Hare R.S., Miller G.H., Shaw K.J.;  
 RT "Analysis of the aac(3)-Vla gene encoding a novel 3-N-  
 RL acetyltransferase."  
 DR Antimicrob. Agents Chemother. 37:2074-2079(1993).  
 FT EMBL: M88012; AAA16193.1; -.  
 SQ SEQUENCE 7 AA; 744 MW; 633862D2C321A030 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
 Db 2 T 2

RESULT 10  
 ID 050556 PRELIMINARY; PRT; 7 AA.  
 AC 050556;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 GN GLTA. (Fragment).  
 OS Actinobacillus actinomycetemcomitans (Haemophilus  
 OC actinomycetemcomitans).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Actinobacillus.  
 OX NCBI\_TaxID=714;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33384;  
 RX MEDLINE=96355846; PubMed=8751884;  
 RA Kolodrubetz D., Spitznagel J. Jr., Wang B., Phillips L.H., Jacobs C.,  
 RT "cis Elements and trans factors are both important in strain-specific  
 RT regulation of the leukotoxin gene in Actinobacillus  
 RL infect. Immun. 64:3451-3460(1996).  
 DR EMBL: U51862; AAB88721.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 832 MW; 6DCB42D767340420 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
 Db 7 A 7

RESULT 11  
 ID 034028 PRELIMINARY; PRT; 7 AA.  
 AC 034028;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 GN Catechol-2,3-dioxygenase (Fragment).  
 OS Sphingomonas chungbukensis.  
 OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;  
 OC Sphingomonas  
 OX NCBI\_TaxID=56193;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DJ77;

RA Kim Y.-C.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U88298; AAB6311.1; -.  
 RN KX  
 FT Dioxigenase.  
 SQ SEQUENCE 7 AA; 868 MW; 71A452D1A699D460 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
 Db 4 T 4

RESULT 12  
 ID 047477 PRELIMINARY; PRT; 7 AA.  
 AC 047477;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)  
 GN tpi. (Fragment).  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE OF 7-7 FROM N.A.  
 RX MEDLINE=85203917; PubMed=3158524;  
 RA Hellenga H.W., Evans P.R.;  
 RT "Nucleotide sequence and high-level expression of the major  
 RL Escherichia coli phosphofructokinase."  
 RL Eur. J. Biochem. 149:363-373(1985).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Evans P.;  
 RL Submitted (OCT-1986) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: X02519; CAA26359.1; -.  
 FT NON\_TER 1  
 SQ SEQUENCE 7 AA; 773 MW; 7A16D33DDDB1DB0 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
 Db 1 A 1

RESULT 13  
 ID 047505 PRELIMINARY; PRT; 7 AA.  
 AC 047505;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 GN MCCA protein.  
 OS Escherichia coli.  
 OG Plasmid pmcc7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9609297; PubMed=8522520;  
 RA Gonzalez-Pastor J.E., San Millan J.L., Castilla M.A., Moreno F.;  
 RT "Structure and organization of plasmid genes required to produce the  
 translation inhibitor microcin C7."

Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1

DB 1 A 1

# RESULT 6

P82541 PRELIMINARY; PRT; 6 AA.  
AC P82541;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DE 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE Chloroplast 30S ribosomal protein S19 beta (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Chloroplast.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;

RN [1]  
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
RC STRAIN=CV, ALMARO; TISSUE=LEAF;  
RX MEDLINE=20435797; PubMed=10874039;  
RA Yamaguchi K., von Knoblauch K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
the small subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 37:28455-28465(2000).  
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -1- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.  
CC -1- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.  
CC -1- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA  
FORM IS THE MINOR BASIC FORM.  
CC -1- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.  
CC InterPro: IPR002222; Ribosomal\_S19.  
DR pfam: PF00203; Ribosomal\_S19; PARTIAL.  
DR PRINTS: PR00975; RIBOSOMAL\_S19; PARTIAL.  
DR PROSITE: PS00323; RIBOSOMAL\_S19; PARTIAL.  
KW Ribosomal protein; Chloroplast; rRNA-binding.  
FT NON\_TER 6  
SQ SEQUENCE 6 AA; 732 MW; 63333735A411C000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 T 1

# RESULT 7

P82182 PRELIMINARY; PRT; 6 AA.  
AC P82182;  
DT 01-JUN-2000 (TREMblrel. 14, Created)  
DT 01-JUN-2000 (TREMblrel. 14, Last sequence update)  
DE 01-MAR-2002 (TREMblrel. 20, Last annotation update)  
DE Chloroplast 50S ribosomal protein L10 gamma (Fragment).  
OS Spinacia oleracea (Spinach).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Caryophyllales; Caryophyllales; Chenopodiaceae; Spinacia.  
OX NCBI\_TaxID=3562;

RN [1]  
RP SEQUENCE.  
RC STRAIN=CV, ALMARO; TISSUE=LEAF;  
RX MEDLINE=20435798; PubMed=10874046;

RA Yamaguchi K., Subramanian A.R.;  
RT "The plastid ribosomal proteins. Identification of all the proteins in  
the 50 S subunit of an organelle ribosome (chloroplast).";  
RL J. Biol. Chem. 275:28466-28462(2000).  
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
CC InterPro: IPR001790; Ribosomal\_L10.  
DR InterPro: IPR002363; Ribosomal\_L10eb.  
DR Pfam: PF00466; Ribosomal\_L10; PARTIAL.  
DR PROSITE: PS01109; RIBOSOMAL\_L10; PARTIAL.  
KW Ribosomal protein; Chloroplast; rRNA-binding.  
FT NON\_TER 6  
SQ SEQUENCE 6 AA; 675 MW; 63218415B05B000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 1 A 1

# RESULT 8

ID 007354 PRELIMINARY; PRT; 7 AA.  
AC 007354;  
DT 01-JUL-1997 (TREMblrel. 04, Created)  
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE MifK (Fragment).  
GN MIFK.  
OS Synechococcus sp. (strain PCC 8801 / RF-1) (Cyanothace PCC 8801).  
OC Bacteria; Cyanobacteria; Chroococcales; Cyanothecae.  
OX NCBI\_TaxID=41431;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RF-1;  
RX MEDLINE=99231861; PubMed=10217509;  
RA Huang T.C., Lin R.F., Chu M.K., Chen H.M.;  
RT "Organization and expression of nitrogen-fixation genes in the aerobic  
nitrogen-fixing unicellular cyanobacterium Synechococcus sp. strain  
RF-1.";  
RL Microbiology 145:743-753(1999).  
DR EMBL: AF003700; AAC35193.1; -.  
FT NON\_TER 1  
SQ SEQUENCE 7 AA; 849 MW; 7412C72A9D5B030 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 7;  
Best Local Similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1  
DB 2 S 2

# RESULT 9

ID 047029 PRELIMINARY; PRT; 7 AA.  
AC 047029;  
DT 01-NOV-1996 (TREMblrel. 01, Created)  
DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)  
DT 01-NOV-1998 (TREMblrel. 08, Last annotation update)  
DE AAD A1 protein (Fragment).  
GN AAD A1.  
OS Enterobacter cloacae.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Enterobacter.  
OX NCBI\_TaxID=550;

DB 1 N 1

## RESULT 2

P83073 PRELIMINARY: PRT: 5 AA.  
 AC P83073: 01-OCT-2001 (TREMBlrel. 18, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
 DE 88 kDa protein (Fragment).  
 OS Bacillus cereus.  
 CC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;  
 CC Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1396;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=NCIMB 11796;  
 RA Browne N., Dows B.C.A.;  
 RL Submitted (JUL-2001) to the SWISS-PROT data bank.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA; 623 MW; 6B01AAA36F00000 CRC64;

Query Match 0.0%; Score 0; DB 2; Length 5;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1

DB 4 T 4

## RESULT 3

O99007 PRELIMINARY: PRT: 5 AA.  
 ID O99007:  
 AC O99007:  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBlrel. 08, Last annotation update)  
 DE Alpha-amylase (EC 3.2.1.1) (Fragment).  
 GN AMY1.  
 OS Hordeum vulgare (Barley).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;  
 CC Trilicaceae; Hordeum.  
 OX NCBI\_TaxID=4513;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HIMALAYA; TISSUE=ALBURNUM LAYER;  
 RX MEDLINE=91329704; PubMed=1831055;  
 RA Jacobsen J.V., Close T.J.;  
 RT "Control of transient expression of chimeric genes by gibberellin  
 RT acid and abscisic acid in protoplasts prepared from mature barley  
 RT aleurone layers.";  
 RL Plant Mol. Biol. 16:713-721(1991).  
 CC -1- CATALYTIC ACTIVITY: ENDOHYDROLYSIS OF 1,4-ALPHA-GLUCOSIDIC  
 CC LINKAGES IN OLIGOSACCHARIDES AND POLYSACCHARIDES.  
 CC -1- COFACTOR: BINDS A CALCIUM ION REQUIRED FOR ITS ACTIVITY.  
 CC -1- MISCELLANEOUS: THERE ARE AT LEAST 4 TYPES OF ALPHA-AMYLASE IN  
 CC BARLEY.  
 DR EMBL: X54643; CAA38455.1;  
 KM Hydroxylase; Glycosidase; Carbohydrate metabolism; Seed; Germination;  
 KW Calcium; Multigene family.  
 FT NON\_TER 5  
 SQ SEQUENCE 5 AA; 600 MW; 61E3344DD6F00000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 5;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1

DB 2 A 2

## RESULT 4

P83308 PRELIMINARY: PRT: 5 AA.  
 AC P83308:  
 DT 01-JUN-2002 (TREMBlrel. 21, Created)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)  
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)  
 DE FMRFamide-like neuropeptide (LPLRF-amide).  
 OS Gallus gallus (Chicken).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 CC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE, AND, SYNTHESIS.  
 RC TISSUE=BRAIN;  
 RX PubMed=6137771;  
 RA Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;  
 RT "A novel active pentapeptide from chicken brain identified by  
 RT antibodies to FMRFamide."  
 RL Nature 305:328-330(1983).  
 CC -1- FUNCTION: MAY FUNCTION AS A NEUROTRANSMITTER OR MODULATOR.  
 CC -1- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)  
 CC FAMILY.  
 KW Neuropeptide.  
 SQ SEQUENCE 5 AA; 645 MW; 69D4073767400000 CRC64;

Query Match 0.0%; Score 0; DB 13; Length 5;  
 Best Local Similarity 0.0%; Pred. No. 0;  
 Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 X 1

DB 1 L 1

## RESULT 5

P82181 PRELIMINARY: PRT: 6 AA.  
 ID P82181:  
 AC P82181:  
 DT 01-JUN-2000 (TREMBlrel. 14, Created)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE Chloroplast 50S ribosomal protein L10 beta (Fragment).  
 OS Spinacia oleracea (Spinach).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 CC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE.  
 RC STRAIN=CV ALMARO; TISSUE=LEAF;  
 RX MEDLINE=20435798; PubMed=10874046;  
 RA Yamaguchi K., Subramanian A.R.;  
 RT "The plastid ribosomal proteins. Identification of all the proteins in  
 RT the 50 S subunit of an organelle ribosome (chloroplast).";  
 RL J. Biol. Chem. 275:28466-28482(2000).  
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.  
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.  
 CC -1- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 KDA.  
 CC -1- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.  
 DR InterPro: IPR001790; Ribosomal\_L10.  
 DR InterPro: IPR002363; Ribosomal\_L10deb.  
 DR Pfam: PF00466; Ribosomal\_L10; PARTIAL.  
 DR PROSITE: PS01109; RIBOSOMAL\_L10; PARTIAL.  
 KW Ribosomal protein; Chloroplast; rRNA-binding.  
 FT NON\_TER 6  
 SQ SEQUENCE 6 AA; 675 MW; 6321BA15B05DB000 CRC64;

Query Match 0.0%; Score 0; DB 10; Length 6;

GenCore version 5.1.4-p5\_4578  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 18, 2003, 14:30:01 ; Search time 3.93103 Seconds  
(without alignments)  
366.908 Million cell updates/sec

Title: US-09-877-606-8  
Perfect score: 7  
Sequence: 1 xxxxxxx 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	0	0.0	4	11	008433	008433 rattus norv
2	0	0.0	5	2	P83073	P83073 bacillus ce
3	0	0.0	5	10	Q99007	Q99007 hordeum vul
4	0	0.0	5	13	P83308	P83308 gallus gall
5	0	0.0	6	10	P82181	P82181 splinacia ol
6	0	0.0	6	10	P82541	P82541 splinacia ol
7	0	0.0	6	10	P82182	P82182 splinacia ol
8	0	0.0	7	2	007354	007354 synechococ
9	0	0.0	7	2	Q47029	Q47029 enterobacte
10	0	0.0	7	2	Q50556	Q50556 actinobacil
11	0	0.0	7	2	Q34028	Q34028 sphingomona
12	0	0.0	7	2	Q47477	Q47477 escherichia
13	0	0.0	7	2	Q47505	Q47505 escherichia
14	0	0.0	7	2	P70804	P70804 azotobacter
15	0	0.0	7	2	Q54248	Q54248 streptomyces
16	0	0.0	7	2	P72081	P72081 nocardia la

17	0	0.0	7	4	Q15897	Q15897 homo sapien
18	0	0.0	7	4	Q15903	Q15903 homo sapien
19	0	0.0	7	5	P83274	P83274 macrobrachi
20	0	0.0	7	6	Q28742	Q28742 oryctolagus
21	0	0.0	7	6	P92372	P92372 haynaldia v
22	0	0.0	7	8	P92403	P92403 lophopyrum
23	0	0.0	7	8	P92425	P92425 pseudoroegn
24	0	0.0	7	8	P92387	P92387 henaradia p
25	0	0.0	7	8	P92427	P92427 peridictyon
26	0	0.0	7	8	P92390	P92390 heteranthe
27	0	0.0	7	8	P92226	P92226 crithopsis
28	0	0.0	7	8	P92214	P92214 amblyopyrum
29	0	0.0	7	8	P92430	P92430 aegilops ta
30	0	0.0	7	8	P92221	P92221 bromus iner
31	0	0.0	7	8	P92442	P92442 taeniathe
32	0	0.0	7	8	P92381	P92381 hordeum bra
33	0	0.0	7	8	P92393	P92393 hordeum vul
34	0	0.0	7	8	P92218	P92218 australopyr
35	0	0.0	7	8	P92440	P92440 thiodopyrum
36	0	0.0	7	8	P92210	P92210 agropyron c
37	0	0.0	7	8	Q99182	Q99182 gnatholebia
38	0	0.0	7	8	Q95945	Q95945 saccharomyc
39	0	0.0	7	8	Q98866	Q98866 splinacia ol
40	0	0.0	7	8	P92421	P92421 psathyrosta
41	0	0.0	7	8	P92385	P92385 hordeum mar
42	0	0.0	7	10	Q49223	Q49223 glycine max
43	0	0.0	7	10	Q95583	Q95583 arabidopsis
44	0	0.0	7	10	P93233	P93233 lycopersico
45	0	0.0	7	10	P82445	P82445 nicotiana t

## ALIGNMENTS

### RESULT 1

ID 008433 PRELIMINARY; PRT; 4 AA.  
AC 008433;  
DT 01-NOV-1996 (TREMUREL, 01, Created)  
DT 01-NOV-1996 (TREMUREL, 01, Last sequence update)  
DT 01-JAN-1999 (TREMUREL, 09, Last annotation update)  
DE UDP-glucuronosyltransferase, microsomal (EC 2.4.1.17) (UDPGR)  
DE (Fragment).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GUNN;  
RX MEDLINE=91282758; PubMed=1840486;  
RA Sato H., Aono S., Kashiwata S., Kotani O.;  
RT "Genetic defect of bilirubin UDP-glucuronosyltransferase in the  
RL hyperbilirubinemic Gunn rat.";  
RL Biochem. Biophys. Res. Commun. 177:1161-1164(1991).  
CC -1- FUNCTION: UDPGR IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND  
CC SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND  
CC ENDOGENOUS COMPOUNDS.  
CC -1- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR - UDP + ACCEPTOR  
CC BETA-D-GLUCURONOSIDE.  
CC -1- SUBCELLULAR LOCATION: MICROSOME.  
CC EMBL: S38636; AAB19259.1; -  
KW Transferase; Glycosyltransferase; Microsome; Multigene family.  
FT NON\_TER 1 1  
FT NON\_TER 4 4  
SQ SEQUENCE 4 AA: 473 MW; 633732C420000000 CRC64;

Query Match 0.0%; Score 0; DB 11; Length 4;  
Best local similarity 0.0%; Pred. No. 0;  
Matches 0; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 x 1